Cloning and characterisation of genes expressed by the third stage larvae of the parasitic nematode *Brugia malayi*

A thesis submitted for the degree of Doctor of Philosophy of the University of Edinburgh

By

William Francis Gregory

Institute of Cell, Animal and Population Biology
University of Edinburgh
Ashworth Laboratories
Kings Buildings
West Mains Road
Edinburgh EH9 3JT

June 2000

ABSTRACT

Human lymphatic filariae currently infect around 120 million people. Larval stages are transmitted by vector mosquitoes and rapidly colonise and develop in the lymphatic system where they live for 5 to 10 years. During their development, larvae present a series of stage-specific antigens to the host's immune system. Successful immunisation regimes in animal models have indicated that antigens presented early on infection are particularly effective in inducing protective immune responses. Epidemiological studies have also suggested that infected individual control their infections by developing immunity to larvae, thereby preventing new infections

Reverse transcription polymerase chain reaction (RT-PCR) of RNA from the infective larval stages of *Brugia malayi* with the nematode-specific 5′ spliced leader sequence and oligo d(T) has been used to identify genes expressed during early larval stages. Products resolved on agarose gels show a number of prominent bands and cloning of these bands has revealed a total of 14 genes. Of these four genes have been analysed in detail.

One highly stage-specific transcript *Bm-cpi*-1, a cystatin-type cysteine proteinase has been characterised in parallel to a second, constitutively expressed, cystatin, *Bm-cpi*-2, identified by EST sequencing. The two inhibitors have been functionally expressed in *E. coli* and have distinct inhibitory specificities. In addition, both CPIs have been localised to the parasite surface and found in parasite secretions.

An abundant transcript, *Bm-alt*-1, is a member of a large family of genes found in nematodes but lacks clear homologues outside the nematode phylum. Analysis of Expressed Sequence Tags (ESTs) deposited in dbEST has identified a total 10 family members in *B. malayi.* Genomic structures of the two most abundant family members have revealed conservation in the position of introns. However there is considerable sequence divergence within introns, with one gene containing distinct repeat units within its introns. In addition, heterogeneity is seen within this due to variation in the number of repeat units.

Two further genes have been characterised. One is a homologue of the human histamine-releasing factor, a constitutively expressed, cytokine-like molecule of interest as a potential modulator of antiparasite responses. A second is a member of a family of proteins of unknown function, rich in glycine and tyrosine residues, which may form part of the nematode cuticle

The discovery of these abundant genes provides candidates for future analysis of immune evasion by filariae and identifies potential targets for vaccination or chemotherapy.

DECLARATION

This declaration is to state that the contents of this thesis have been composed solely by myself, and that the work described is entirely my own.

ACKNOWLEDGMENTS

I would like to thank the many people who made my project not only a great learning experience but also more pleasurable than any PhD student should expect!

Thank you to my supervisor Rick Maizels for direction, encouragement and friendship throughout my (long) stay in his lab. Thanks for giving me the chance. I hope we will always stay in contact. Thanks also to Mark Blaxter for helping me enter the world of molecular biology and genomics and to Judi Allen for guidance and friendship.

To all the members of Ricks Lab: Natalia Gomez-Escobar, Janice Murray, Kevin Tetteh, Alex Loukas, Andrew MacDonald, Xingxing Zang, Yvonne Harcus, Agnes Kurniawan-Atmadja and Franco Falcone. A great team. Thanks for all your willing help and intellectual input... and of course all the parties.

To David Guiliano for help in almost all aspects of my project, especially once the great influx of EST sequences had begun. Thanks to P'ng Loke for his interest in my project, weekend chats over coffee and great ideas.

CONTENTS

CHAPTER 1

INTRODUCTION

Lymphatic filariasis is a chronic, debilitating and often disfiguring disease that affects about 100 million people living in the rural and slum areas of many tropical countries (Michael et al., 1996; World Health Organization, 1995). The causative agents in man are three species of filarial worms, *Wuchereria bancrofti, Brugia malayi* and *B. timori*. These are multicellular, parasitic nematodes with a complex life cycle that includes an obligatory maturation period within a vector mosquito. The infection is widespread throughout the tropical and subtropical areas of Asia, Africa, the Western Pacific and some parts of the Americas.

1.1 Life Cycle

In common with all nematodes, filarial worms progress through five developmental stages during their life cycle (designated L1 to L5). Four of these stages are delineated by moults - the shedding of a mutilayered collagenous exoskeleton termed the cuticle. The mature (adult or L5) stages differ in size. They are thread-like in appearance: the male measures 22 mm in length and 0.1 mm in diameter while the female is considerably longer measuring 48 mm in length and 0.15 mm in diameter. After mating, the female constantly produces thousands of embryonic first stage larvae, the microfilariae (Mf). Adult female parasites remain reproductively active for 5 - 10 years in the lumina of the host lymphatic system (Vanamail et al., 1989). Mf are enclosed by a acellular sheath, a complex structure which is a remnant of the vitelline membrane formed during in utero development. Mf migrate to the blood stream where they are taken up with the blood ingested by a blood feeding vector. In the mosquito midgut they shed their sheaths and penetrate the gut wall to reach the haemocoele through which they migrate to the thoracic flight muscles. Lymphatic filarial larvae at this stage of development are almost unique amongst nematodes as they occupy an intracellular niche within the

large cells of the thoracic muscle. Here they develop into thicker, shorter "sausage" forms. This is followed by a programmed set of cell divisions that result in the differentiation of the oesophagus, intestine and rectum (Schacher, 1962b). The L1 to L2 moult takes place between days 6 and 10 post-infection depending upon environmental conditions. The L2 to L3 moult occurs between days 8 and 11 post-infection by which time the gut appears longer and more developed and the genital primordium is formed. The L3 larvae measure 1.5 mm in length and 0.02 mm in diameter (Schacher, 1962b).

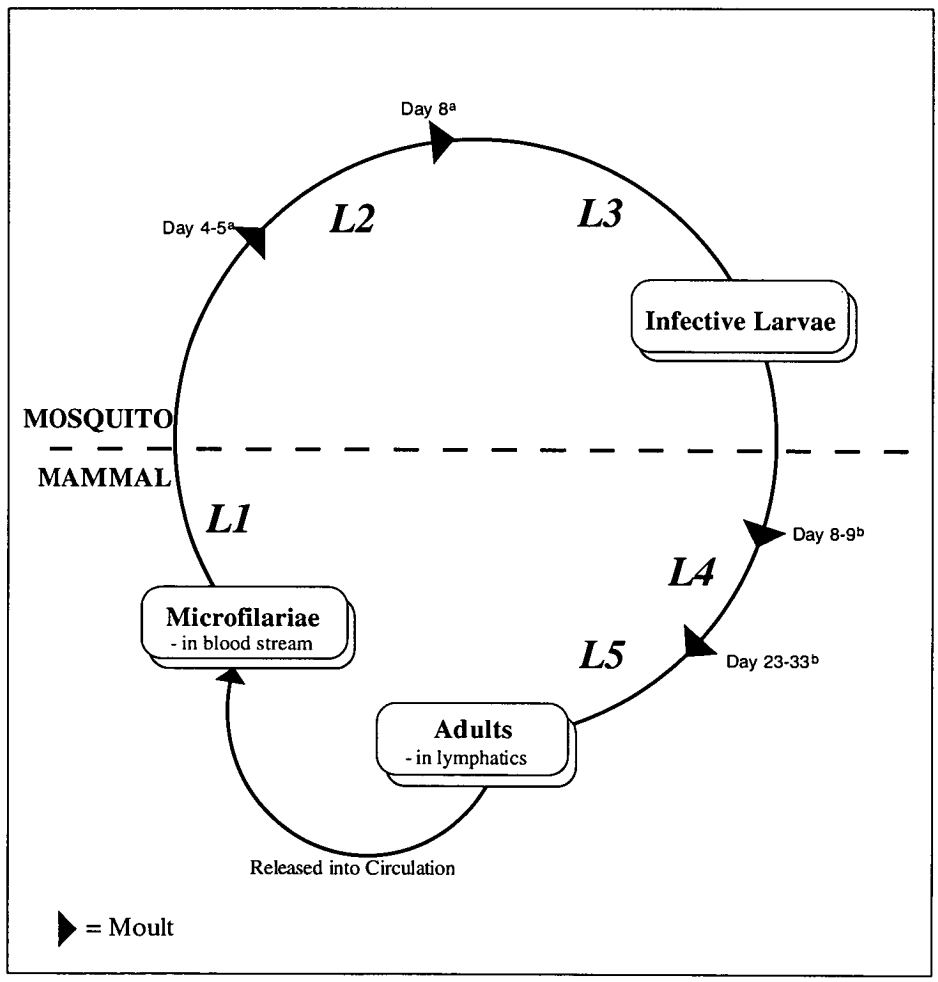
Considerable attrition occurs during the vector stages of development. Infection levels varying depending upon the particular filariae-vector combination. Even in a susceptible mosquito host, such as *Aedes aegypti*, up to 80% of Mf fail to develop to infective larvae. Parasites are killed through a combination of physical incompatibility and an active immune response against the parasites (Denham and McGreevy, 1977). Two loci are known to confer resistance to *B. malayi* in *Ae. aegypti*. One is the originally described f^m gene, which affects susceptibility to filarial worms inhabiting the thoracic musculature (MacDonald and Ramachandran, 1965). The other locus exhibits an additive effect on susceptibility. Whereas the products of these genes probably act on the parasite within the thoracic muscles, physical factors act in the fore and midgut to prevent infection. Some species have pharyngeal armatures capable of damaging Mf as they are ingested. Clotting of the blood meal is also important as Mf are trapped in the gut of mosquitoes which show rapid coagulation. In the midgut other mechanisms have also been suggested, including damage by digestive enzymes, limitation of migration by the peritrophic membrane and antifilarial toxins (Denham and McGreevy, 1977). One response of the mosquito to filarial infection seen in some, but not all, vector/filariae combinations is melanotic encapsulation (Nayar et al., 1993; Oothuman et al., 1974; Zhao et al., 1995). This response kills the parasites in capsules composed of haemocytes and/or polyphenolic materials and can take place in the haemocoele or the thoracic muscles (Chikilian et al., 1994).

Transmission to a mammalian host is possible soon after the second moult. At this point the larvae migrate from the thoracic muscles and distribute themselves

throughout the body cavities of the mosquito with a strong preference for the head and labium (Schacher, 1962b). Development is then arrested until transmission to the final host is accomplished. During feeding the larvae burst out of the proboscis and remain in a pool of haemolymph around the wound. After the mosquito has withdrawn its proboscis the larvae enter the new host through the lesion left by the bite (McGreevy et al., 1974). In cats it has been shown that larval migration to the nearest lymph node is accomplished within a few hours of infection (Ewert and El Bihari, 1971; Ewert and Ho, 1967). Later they migrate to the afferent lymphatic where they stay for the rest of their lives. The mechanisms by which the larvae seek, recognise and colonise the lymphatics are unknown but presumably involve interactions with host factors. Proteolytic enzymes such as collagenase may be involved in these processes (Petralanda et al., 1986). In the definitive host the L3 re-initiates its developmental program that culminates in a moult to fourth stage larvae at around day nine post-infection. Newly formed L4 larvae are 3 to 6 mm in length. A final L4 to adult moult takes place 3 to 4 weeks post-infection. Adult male worms are distinctly smaller than the females (around 4 cm, compared with 8 – 10 cm)

After transmission, filariae exit from a state of developmental arrest. The early period of infection is characterised by an intense period of development and adaptation to a new, radically different environment. Within the first 2 weeks of infection 2 of the 4 moults are completed (Fig 1). Previous studies have shown that the surface of the early infective forms of several nematode species is a dynamic structure. For the intestinal parasites *Trichenella spiralis* and *Toxocara canis* it has been demonstrates that the infective larvae release a significant amount of their surface molecules during the first few hours of development (Maizels et al., 1984; Philipp et al., 1980). The dynamic nature of the surface has also been observed for filarial larvae. After labelling surface components with ^{125}I, *B. malayi* larvae lost over 90% of labelled molecules after eight days of jird infection (Marshall and Howells, 1986). Radiolabelling has also been used to demonstrate the variety of highly stage specific proteins on the surface of *B. malayi* (Storey and Philipp, 1992) and *A. viteae* (Apfel et al., 1992) during the first few days of infection. This phenomenon has also been shown for *B. malayi* larvae undergoing L2-L3 development in the vector (Ibrahim et al., 1992). These changes, indicated by

changing patterns of labelled proteins as development proceeds, are largely uncharacterised at the molecular level and few genes corresponding to these proteins have been identified. Some studies have attempted to discover cues for these changes. A 23 kDA molecule was prominently labelled after micro-injecting [^{35}S]methionine into vector blackfly during the later stages of *Onchocerca lienalis* and *O. volvulus*. During subsequent *in vitro* culture shifting temperature from 26°C to 37°C specifically triggered secretion of the molecule (Bianco et al., 1990; Bianco et al., 1995). This molecule has subsequently been identified as a member of the *Onchocerca* family of abundant larval transcripts (*alt*) (Yang Wu and Ted Bianco, University of Liverpool, personal communication), homologous to the *B. malayi alt* genes described in chapter 5.



[a] Data from *Brugia pahangi* infections in *Anopheles* spp, Schacher (1962a)
[b] Data from *Brugia pahangi* infections in the cat, Schacher (1962b)

Figure1.1. *The Life Cycle of Lymphatic Filarial Parasites*

1.2 Human Lymphatic Filariasis

Filarial worms are noted for their narrow host specificity. The superfamily Filariodea contains 535 species infecting almost all main orders of vertebrates but only eight are known to infect man (Sasa, 1976). Of the filariae that infect humans, only the lymphatic filariae (*Wuchereria bancrofti*, *Brugia malayi* and *B. timori*) and the cutaneous dwelling filariae, *Onchocerca volvulus,* are considered of major public health importance.

More than 1.1 thousand million people (20% of the world's population) now live in areas where they are at risk of infection with lymphatic filarial parasites (Ottesen et al., 1997) and a minimum of 120 million people are currently infected (Michael et al., 1996). Despite causing little mortality, lymphatic filariasis has been identified as the leading cause of long-term disability (World Health Organization, 1995) leading to huge economic losses. In India, conservative estimates indicate that the indirect economic cost of lymphatic filariasis is US$ 1.5 billion per year (Ottesen et al., 1997).

A total of 44 million people suffer from pathological effects of the infection (Ottesen et al., 1997). Clinical manifestations in all three human lymphatic filariae are very similar and include painful attacks of lymphadenitis and chronic obstructive lesions such as lymphoedema, and elephantiasis (Partono, 1987). These obstructive lesions are most probably caused by immune-mediated destruction of resident adults and migrating larvae in the lymphatic system (Maizels and Lawrence, 1991; Ottesen, 1992).

Not all people in an endemic area, however, present symptoms. One of the most intriguing aspects of filarial infections is the extremely broad spectrum of clinical presentation found among individuals in endemic regions (Ottesen, 1992). At one extreme there are individuals with no clinical manifestations or indication of infection at all, although epidemiological studies demonstrate that exposure to

infective larvae is equivalent to those individuals with clinically expressed filariasis. It is difficult to ensure that this group of people, termed endemic normals, consist only of those truly immune and parasite-free, and to exclude those who harbour a sub-clinical infection, as rigorous techniques to determine parasitological status, such as lymphosonography, are not in widespread use. Sera endemic normals contain high levels of antifilarial antibody and active cellular responses to the parasite can be demonstrated (Maizels et al., 1995). At the other extreme, are the group of people exhibiting clinical symptoms ranging from acute (recurrent episodes of filarial fever accompanied by inflammation of lymphatic tissue) to chronic (obstructive disease including elephantiasis of the limbs and other regions and hydrocoeles).

Two other categories can be distinguished in endemic areas. The first are those who are entirely asymptomatic but carry detectable levels of Mf in their peripheral blood. The final group suffers from tropical pulmonary eosinophilia, an occult form of filariasis characterised by extremely high levels of IgE, eosinophilia and asthmatic attacks. Circulating Mf are not found in these individuals but they may sometimes be found in biopsies of lung and other tissues.

Epidemiological evidence suggests that many residents of endemic areas eventually progress through this clinical spectrum from asymptomatic amicrofilaraemics to chronic obstructive disease (Bundy et al., 1991; Day et al., 1991b). In most endemic areas, the prevalence of acute filarial disease increases between birth and early childhood, and reaches a plateau in older age groups. Acute symptoms precede the onset of chronic obstructive disease. Hydrocoeles and elephantiasis are rarely found in adolescents but become increasingly common with prolonged residence in an endemic area and some studies have shown a temporal correlation of the onset of chronic pathology with the disappearance of Mf from the circulation. However conversion to chronic disease is not inevitable, some long term studies have described communities in which 80% of microfilaraemics remain disease-free for 16 years (Meyrowitsch et al., 1995). The possibility that lymphatic lesions are the products of the immune-mediated destruction of resident adult worm and migrating larvae presents a formidable task for immunoparasitologist in the

development and implementation of vaccines against filarial worms. Thus a key question is whether it is possible to stimulate immunity without accelerating the progression of individuals from the asymptomatic to the pathologic groups.

1.3 Animal Models

A significant obstacle to understanding the biology and immunology of the most pathogenic and widespread filarial worms, *W. bancrofti* and *O volvulus,* is that they cannot be maintained in laboratory animals. Infections in monkeys (Misra et al., 1997) and chimpanzees (Greene, 1987) have been reported but infection levels are poor, and these animals are extremely difficult to obtain due to their protected status.

Fortunately, *Brugia* spp. is more cosmopolitan. *B. malayi* develops to maturity in jirds rats, hamsters, ferrets, dogs and cats with varying degrees of success and the natural filarial worm of the cat, *Brugia pahangi*, also infects a similar range of laboratory animals with equal if not greater success. The jird (*Meriones unguiculatus*) is the most susceptible animal for both species with up to 50% of larvae developing to adults. This level is only achieved if the larvae are inoculated into the peritoneal cavity, where adults and Mf remain for the duration of the infection. Other routes of infection result in substantially reduced recovery of adult worms and Mf. For all routes of infection male jirds are markedly more susceptible than their female littermates. Thus intraperitoneal infection of male jirds is now routinely used for the maintenance of *B. malayi* and *B. pahangi* in many laboratories (McCall et al., 1973).

The rat and the multimammate rat (*Mastomys natalensis*) are variably permissive to infection with *Brugia* spp. However, the obvious choice for a convenient, well-defined laboratory host for filarial infection would be the inbred mouse. Unfortunately, numerous attempts at infection of a wide variety of strains have shown that mice are largely resistant to the development of infective larvae of all species of human lymphatic filariae (Ahmed, 1967; Laing et al., 1961). Again most strains are more permissive to *Brugia* spp, but larvae can only survive for

short periods of time and will undergo limited development before the parasites are cleared (Lawrence, 1996). Reports vary considerably in their estimates of how long worms survive. One of the most thorough investigations (Carlow and Phillip, 1987) reported that less than 30% of worms are recovered 10 days after a primary infection and none after day 30. Secondary infections showed considerably increased resistance to infection, with no recovery after 10 days. Implants of adults and Mf parasites can however be maintained in mice for longer periods, allowing the study of parasite factors on host cell populations (Allen et al., 1996; Lawrence, 1996)

Despite a lack of full development, some characteristics of human filariasis can be seen in during infection of mice with *Brugia* sp. Mice can develop strong Th2-cell responses, anergised immune responses and some pathological reactions in response to infection (Lawrence, 1996). In addition, SCID and athymic mice are susceptible to full infection with the development of Mf-producing adults and develop lymphatic pathology that is reversible if adult worms are removed by chemotherapy (Nelson et al., 1991; Vickery et al., 1985). If athymic mice are reconstituted with splenocytes from an intact donor that has been previously primed with infective larvae progressive lymphatic fibrosis that is irreversible is seen which has the appearance of the pathological effects of elephantiasis (Vickery et al., 1991). Thus it seems that pathology in mice has both immune, and non-immune components.

The course of infection of cats infected with *B. pahangi*, which parasitises them in the wild, mimic those of humans infected with *B. malayi* and *W. bancrofti* in many important respects (Denham and Fletcher, 1987). Multiple inoculation of larvae representing the kind of exposure experienced by patients living in endemic areas results in transient enlargement of the lymphatics. Most cats develop patent infections but eventually lose their circulating Mf suddenly. These post microfilaremic cats exhibit prominent, but temporary, lymphodema. Although permanent damage leading to elephantiasis is not seen as it is in human infections it is likely that this natural host-parasite relationship displays many features of human infection. Despite this, the *Brugia*-cat model is far from ideal when issues of

availability, cost and lack of suitable reagents to study immune responses are taken into account.

Litomosoides sigmodontis, a natural filarial parasite of the cotton rat, has recently been shown to undergo full development in mice (Petit et al., 1992). The L sigmodontis/mouse model has a number of features useful for the investigation of the host response to filariae. In addition it is a convenient primary screen for drug and vaccine candidates. In the BALB/c mouse subcutaneous infection of is characterised by a initial period during which up to 75% of larvae are killed before penetration of the host lymphatic vessels. Thereafter, the parasite burden is stable for around 60 days (Marechal et al., 1996). Infections in other strains of mice show varying degrees of resistance and unlike resistance induced by irradiated larvae attrition occurs at the L4 and adult stages. In B10D2 mice the recovery rate during the first month is similar to that seen in BALB/c mice but thereafter it declines without producing a patent infection (Petit et al., 1992). The differences in susceptibility in various mouse strains means that this model may have the potential to increase understanding of the immune response to filariae in a way that *Trichuris muris* has for intestinal nematodes (Grencis, 1996) and Leishmania major has for intracellular protozoa (Reiner and Locksley, 1995). It is interesting that in the *Litomosoides* model the pattern of susceptibility (BALB/c susceptible, B10 resistant) is the reverse of the Trichuris model (Grencis, 1996) as it is thought that resistance to nematode infections is defined by the hosts ability to mount a strong Th2 immune response. BALB/c mice are able to mount a strong Th2-type response but are fully susceptible to L. sigmodontis (Marechal et al., 1997).

1.4 Chemotherapy

At present, control of lymphatic filariasis is dependent on repeated anthelminthic treatment and vector control regimens. Until recently, the only anthelminthic used to treat lymphatic filariasis was diethylcarbamazine (DEC). Administration of the drug leads to rapid decreases in circulating Mf and repeated administration is reported to kill a proportion of the adult worm burden. Various treatment regimes have been used usually with higher doses giving rise to side reactions (fever, vomiting, painful lymph nodes and headaches) associated with the

death of Mf (Mackenzie and Kron, 1985). An alternative approach to the administration of DEC is its addition to salt. This allows very low doses to be given through daily use of salt and has been shown to dramatically reduce microfilarial density with no apparent adverse side reactions (Gelband, 1994).

Recent studies have shown the combination of ivermectin and DEC, given on a yearly basis, is superior to either drug given alone, both in parasite reduction and in sustaining the effect for one year (Ismail et al., 1998; Kazura et al., 1993; Moulia-Pelat et al., 1995; Nicolas et al., 1997; Shenoy et al., 1998). DEC and ivermectin are much less effective against adult parasites with studies using ultrasonography to detect adult parasites in situ reporting no effect (Dreyer et al., 1996; Dreyer et al., 1998). This raises the problem of a recurrence of transmission if the drug is not re-administered to the whole community. The possibility of drug resistance also exists if only partial killing of parasites is attained. Another drawback is that DEC is not recommended by the World Health Organization (WHO) for use in areas where lymphatic filariasis exists alongside *Loa Loa* due to the quick killing of their Mf, which leads to dangerous adverse reactions (Ottesen et al., 1997). Administration of these drugs does not prevent the re-establishment of infection and there is no clear evidence that killing an existing worm burden results in a boosting of the immune response to incoming larvae. Treatment is required at least once a year until both adults and Mf are killed and infection is eliminated from the vector population and animal reservoirs.

Many species of filariae harbour intracellular bacteria within the cells of the hypodermis and reproductive tissues, representing a novel target for antifilarial drugs. First seen in electron micrographs they are related to the *Wolbachia* symbionts of vector insects (McLaren, 1975). Sequencing of 16S ribosomal genes from bacteria isolated from different filariae have shown that they are a distinct group inhabiting almost all filariae except the rodent filariae, *A. viteae*, and the deer parasite *Onchocerca flexuosa*. An important finding is that treatment of filarial infections with antibiotics effective against rickettsial bacteria reduces larval development and renders adult worms infertile (Hoerauf et al., 1999). Investigations into whether antibiotics are directly toxic to worms or whether

worms die if their bacteria are eliminated have suggested the latter. Tetracycline is ineffective against *A. viteae*, which lack *Wolbachia* and loss of bacteria is seen long before worm death in *L. sigmodontis* (Hoerauf et al., 1999). This raises the possibility that filariae rely on some of the genome of *Wolbachia* for establishment or maintenance in their mammalian host. It is however possible that death of *Wolbachia* within worms releases toxins. Nethertheless treatment of the bacteria would be a useful addition to available therapies.

1.5 Vaccination

A major goal of current research on parasitic diseases is vaccine development. Mass vaccination programs offer the opportunity of long term immunity, ideally, without the need for boosting. Clearly the first need is to identify which of the many thousand potential antigens encoded within the filarial genome will be useful in a vaccine. Equally important is the need for a clear understanding of the immune mechanisms that drive parasite clearance as distinct from the mechanisms that induce pathology.

In endemic areas evidence does exit for a degree of protective immunity in long term residents of infective areas. By far the largest group within an endemic area are those people who have no clinical symptom of infection and are amicrofilaraemic. Although a proportion of these individuals will have infection levels undetectable by current methods some will undoubtedly have effective immunity. Moreover, by grouping infected individuals by age it has been shown that accumulation of adult worms tapers off at around 20 years of age (Vanamail et al., 1989) and that this coincides with the presence of antibodies to the surface of the infective larvae (Day et al., 1991a). Taken together these data indicates the existence of concomitant immunity similar to that found in schistosomiasis (Woolhouse and Hagan, 1999).

Animal models have provided the strongest evidence that protective immunity can be induced and that the larval stages are the most effective target. Substantial levels of immunity can be generated by irradiated L3s (Oothuman et al.,

1979; Storey and Al Mukhtar, 1982; Wong et al., 1969; Yates and Higashi, 1985), trickle infections (Denham et al., 1992; Lucius et al., 1986), drug abbreviated infections (Chusattayond and Denham, 1984; Grieve et al., 1988) and secreted larval antigens (Lucius et al., 1991).

The most widely used and successful of these techniques is inoculation of irradiated larvae. The first irradiation attenuated larval vaccines to be developed were against the cattle lungworm *Dictyocaulus* (Peacock and Poynter, 1980) and *D.filaria* of sheep (Jovanovic et al., 1965). Both vaccines proved to be very successful and are still widely used in the veterinary field.

The exact effect of radiation on the biology of the larval parasites and the immune mechanisms involved in resistance against infective larvae are unknown. Depending upon the dose, irradiation will retard parasite maturation and migration to the preferred site of development within the host. Moreover, natural infections have been reported to produce more lymphatic pathology than infections with irradiated larvae (Oothuman et al., 1979). It has been suggested that the basic principle underlying an attenuated vaccine may be to stimulate immune responses towards the early stages of its development in the mammalian host (Denham and McGreevy, 1977). Thus when the immunised host is re-infected the immune responses could destroy the larvae, preventing infection. It is known that after exposure to ionising radiation, larvae spend longer periods in the L3 and L4 stages and thus larval antigens expressed only in the early stages of development will be presented to the host for longer periods. Importantly, a general observation has been that dead worms or somatic extracts provide modest or no protection (Abraham et al., 1988; Carlow and Phillip, 1987).

There is some evidence that other life cycle preparations can be effective at killing parasites. Immunisation with soluble microfilarial extract is effective at reducing the numbers of jirds reaching patency and also reduces adult worm burden (Kazura et al., 1986). A 75 kDa surface-associated chitinase is the target of protective monoclonal antibodies (Fuhrman et al., 1992), and vaccination reduces microfilaraemia without affecting adult worm numbers (Wang et al., 1997).

Species	Vaccine	Host Species	Challenge	Protection	Reference
B. malayi	Irradiated L3	Jird	L3	44–91%	(Yates and Higashi, 1985)
	Irradiated L3	Mouse	L3	95.5%	(Hayashi et al., 1984)
	Dead L3	Mouse	L3	46%	(Carlow and Phillip, 1987)
	Aqueous L3 extract	Mouse	L3	76%	(Carlow and Phillip, 1987)
	Aqueous Mf extract	Mouse	L3	56%	(Kazura et al., 1986)
B. pahangi	Irradiated L3	Cat	L3	80%	(Oothuman et al., 1979)
	Irradiated L3	Jird	L3	39-76%	(Chusattayanond and Denham, 1986)
	Irradiated L3	Mouse	L3	94%	(Hayashi et al., 1984)
A. viteae	Irradiated L3	Jird	L3	60-90%	(Lucius et al., 1991)
	Exported antigens	Jird	L3	67.4%	(Lucius et al., 1991)
D. immitis	Irradiated L3	Dog	L3	88%	(Wong et al., 1969)
	Albendozole-cured infection	Dog	L3	97%	(Grieve et al., 1988)
L. sigmodoites	Irradiated L3	Mouse	L3	50-57%	(Le Goff et al., 1997)

Table 1.1: Non-defined vaccine preparations used to protect against filarial infections in animal models.

Although microfilarial antigens have the potential to block transmission, the effect of immunising actively infected individuals against Mf is unknown. Pathological reactions to dying Mf may result. Taken together the infective larval stage is the most attractive stage to attack with a vaccine since the vulnerability of this stage has been demonstrated and pathological reactions are likely to be less severe compared with other stages.

1.6.1 Identification of protective antigens

Work to date on selecting molecules for a filarial vaccine has largely been concerned with the identification of protective antigens without a full understanding of what determines a protective immune response. Most effort has been concentrated on identifying differences in antibody recognition in infections displaying different clinical and parasitological manifestations. This, however, has

several limitations. At present there is little evidence for antibody-mediated protection against filarial infection. Infections in mice have shown that T cells alone can mediate clearance of an infection (Suswillo et al., 1981), indicating that antigens capable of stimulating good T cell responses would also be important. If antibodies are important in protection it may be that the effector mechanisms induced by the antibody are important, rather than antibody per se. Another problem is defining a response that is linked with protection in patient sera. The antibody profiles of individuals will represent the recent history of the infection including responses to dead larval parasites and Mf. Thus the antibody profile of an infected individual will reflect the antigenicity of each molecule, which is unconnected to its potential as an immunogen. Thus many studies have identified highly immunogenic paramyosin (Li et al., 1991) and heat shock proteins (Selkirk et al., 1987), tropomyosin, (Hartmann et al., 1997a), chitinase (Adam et al., 1996) and myosin (Wang et al., 1997). These responses may be evoked as a consequence of parasite death but not be a cause of it. Recently, vaccination of jirds with highly immunogenic proteins (heat shock protein, collagen and myosin) showed that they failed to induce protective responses (Peralta et al., 1999). Differential screening by western blot identified a 43 kDa L3 chitinase recognised by all patients rigorously selected as putatively immune (endemic normals) but infrequently but patients harbouring Mf (Freedman et al., 1989; Raghavan et al., 1994).

Another approach has been to define antigens from larval stages without selecting for antibody reactivity. The success of live, irradiated vaccines and excretory/secretory (ES) antigens (Lucius et al., 1991) contrasted with the relative inefficacy of dead preparations suggest that key antigens are expressed following infection. Careful implantation of various stages of post parasitic larvae has identified larvae undergoing the L3 – L4 moult as the most potent source of protective antigens (Eisenbeiss et al., 1994). Highly stage-specific antigens have also been defined (Bianco et al., 1990; Bianco et al., 1995; Frank and Grieve, 1991; Frank and Grieve, 1996; Frank et al., 1996; Scott et al., 1990; Storey and Philipp, 1992). These studies not only offer the chance of identifying potential vaccine candidates but offer insights into the biology of the parasite during this stage of development. Efforts have begun to identify genes encoding proteins specific for

the infective larval stage (Blaxter et al., 1996; Gregory et al., 1997; Martin et al., 1996; Scott et al., 1995; Yenbutr and Scott, 1995).

1.6.2 Surface antigens

By analogy with microbial infections where surface antigens are the targets of protective immune mechanisms, antigens expressed on the surface of filariae are considered likely vaccine candidates. In addition information about the biology of the parasite can be gained by identifying and characterising these proteins. A mouse monoclonal antibody against a chitinase present on the surface of fully mature Mf is capable of reducing infections in vector mosquitoes and Mf levels in infected gerbils. The homology with chitinases suggests a role in the penetration of the mosquito midgut. The surface-associated proteins of adult parasites originally visualised by surface-specific labelling techniques are now being identified. The dominant protein is gp29/*Bm-gpx* (Cookson et al., 1992), a secreted glutathione

Gene Name	Identification	Species	Vaccination model	Stage specificity	Reference	Protection (stage)
MOv 2	Retinol-binding protein	*O. volvulus*	*A. viteae/jird*	All stages	(Jenkins et al., 1996)	45.5%*(adult) 89% (Mf)
Mov14	Tropomyosin	*O. volvulus*	*O. liensis/mouse*	All stages	(Taylor et al., 1996)	48 - 62% (Mf)
			A. viteae/jird		(Taylor et al., 1996)	46% (adults)
OvB20	Novel	*O. volvulus*	*A. viteae/jird*	Mf, L3	(Taylor et al., 1995)	54.5%*(adult) 97% (Mf)
			O. lienalis/mouse		(Taylor et al., 1995)	NS
AvL3-1	LIM domain protein	*A. viteae*	*A. viteae/jird*	ND	(Oberlander et al., 1995)	NS
OvG15	Heat shock protein 70	*O. volvulus*	*B. malayi/jird*	?	(Peralta et al., 1999)	NS
K11	Myosin	*O. volvulus*	*B. malayi/jird*	?	(Peralta et al., 1999)	NS
AP2	α1-type IV collagen	*B. malayi*	*B. malayi/jird*	?	(Peralta et al., 1999)	NS
F7R2	Chitinase	*B. malayi*	*B. malayi/jird*	Mf	(Wang et al., 1997)	† (Mf)
SXP-1	Novel	*B. malayi*	*B. malayi/jird*	?	(Wang et al., 1997)	† (Mf)
BM5	Paramyosin	*B. malayi*	*B. malayi/jird*	All stages	(Li et al., 1993)	43 (Mf)
62 kDa antigen	Aspartyl tRNA synthetase	*B. malayi*	*B. malayi Mf/mouse*	Mf, adults	(Kazura et al., 1990)	40 – 60%
Ov-ALT-1	Novel	*O. volvulus*	*O. volvulus/mouse*	All stages	(Joeseph et al., 1998)	39%

*Average of 2 experiments
NS, not significant
ND, not determined
†, significant but % reduction not given in the text

Table 1.2. Recombinant antigens used to protect against a challenge of filarial parasites.

peroxidase with a preference for lipid substrates (Tang et al., 1995). Two other surface-associated antioxidants, superoxide dismutase (Ou et al., 1995) and thioredoxin peroxidase (Ghosh et al., 1998). Another major surface protein present in preparations for all life cycles stages is p15/400, nematode polyprotein antigen, a lipid binding protein possibly involved in scavenging of host lipids (Kennedy et al., 1995).

1.6.3 Excretory/secretory products.

Lymphatic-dwelling filarial parasites release a wide variety of antigens when maintained in culture (Kaushal et al., 1982; Kwan-Lim et al., 1998; Maizels et al., 1986; Maizels et al., 1985), including anti-oxidants (Cookson et al., 1992; Tang et al., 1994), retinol binding protein (Selkirk et al., 1993), and acetylcholine esterase (Rathaur et al., 1987) macrophage migration inhibitory factor (MIF) (Pastrana et al., 1998) and protease inhibitors (Yenbutr and Scott, 1995). Attention has focused on released molecules with the assumption that accessibility to the effector arms of the immune system is a prerequisite for effective vaccine candidates. In addition secreted components potentially include immune modulators and molecules diagnostic of active infection.

1.7 The filarial genome

The filarial nuclear genome is similar in size to that of *C. elegans*, the first multicellular organism to be sequenced in full (The *C. elegans* Sequencing Consortium, 1998). There are predicted to be around 19,000 proteins coding genes in *C. elegans* and a similar number is expected in filariae. In common with *C. elegans,* and perhaps universal amongst nematodes, is the *trans*-splicing of most mRNAs (Blaxter and Liu, 1996).

Trans-splicing involves the addition of short leader sequences onto the 5′ ends of mRNA (Nilsen, 1993). In trypanosomatidae all nuclear mRNAs are *trans*-spliced to a 35-96 nt sequence at their 5′end donated from a 137 nt precursor. In

nematodes some, but not all, mRNAs receive a 22 nt *trans*-spliced sequence (SL-1) from a 100 nt donor (Blaxter and Liu, 1996). The significance of this sequence is not yet known but it is estimated that more than 70% *C. elegans* transcripts receive this sequence during maturation of the primary mRNA transcript.

C. elegans is unusual among eukaryotes in that a significant number of its genes are arranged in polycistronic pre-RNAs derived from co-transcribed gene clusters, or operons (Spieth et al., 1993; Zorio et al., 1994). A further, distinct, *trans*-spliced sequence, SL2, has also been found on the 5′ end of a much more limited set of mRNAs and at present it appears that this splicing event is associated with downstream genes in operons (Spieth et al., 1993). Whereas the first gene in an operon receives SL1 or is not *trans*-spliced at all, genes located downstream either receive SL2 or a mixture of SL1 and SL2 (Evans et al., 1997; Zorio et al., 1994). In *C. elegans* other *trans*-spliced sequences (SL1 – 5), more closely related to SL-2 than SL-1, have been described. (Ross et al., 1995).

A common feature of operons in *C. elegans* is that each gene is about 100 bp downstream from another gene with the same 5′-3′orientation. Searching two megabases of *C. elegans* genomic sequence has revealed that perhaps 25% of genes are organised in operons (Zorio et al., 1994).

It is likely that SL-1 and SL-2 do not solely function within operons as there are many *trans*-spliced genes not organised in such a way. The sequence probably acts as a good translational initiation site, although other functions are possible (nuclear export or protection of mRNA from degradation (Maroney et al., 1995), for instance)

Although the function of the splice leader sequence is not fully understood they can be conveniently used to isolate the 5′ end of cDNAs or to construct libraries (or pools of 1st strand cDNA) (Blaxter et al., 1996; Gems et al., 1994; Gems and Maizels, 1996; Gregory et al., 1997; Martin et al., 1996; Scott et al., 1995; Yenbutr and Scott, 1995).

Filarial nematodes possess six chromosomes (5 autosomes and a XY sex determination pair). A large portion of the 100 Mbp genome is repetitive DNA. Approximately 10% of the *B. malayi* genome is made up of a single repetitive element, the *Hha*I repeat (around 30000 copies/genome) (McReynolds et al., 1986). This 322 nucleotide repeat is arranged in tandemly repeated arrays and is found in similar copy numbers in other species of *Brugia* but not in *W. bancrofti* (Xie et al., 1994). Overall the genome has a high AT content of 71% (Rothstein et al., 1988).

Up to 1994 a small set of around 60 *Brugia* gene sequences had been submitted to the public databases, largely identified through immunoscreening of adult cDNA libraries. Given an estimated complement of around 15 – 20000 genes encoded within the genome a radical departure in methodologies for gene discovery was needed. In the last few years, the Filarial Genome Project has been initiated to facilitate parasite gene discovery. A major part of the project has been the generation of a series of cDNA libraries representing all stages of the life cycle of *B. malayi*. Genes are then identified by the Expressed Sequence Tag (EST) approach. Clones are selected randomly from cDNA libraries, and single pass DNA sequencing reactions are performed to provide a 100 – 500 nucleotide sequence "tag" for each selected clone. As each gene is represented in a cDNA library at approximately the abundance of its mRNA, a record of gene expression through the life cycle can be generated (Blaxter, 1995). This effort is ongoing and there are currently around 20,000 ESTs deposited in the public database, dbEST, defining around 6,500 genes. If the total number of genes in *B. malayi* is similar to that predicted from the entire genome of *C. elegans* then around one third of *B. malayi* genes have already been identified. Although the entire sequence of the *C. elegans* genome is now available, 38% of *B. malayi* ESTs are novel, having no matches to sequences in the public databases (Williams et al., 1999). These potentially represent a collection of filarial-specific genes and are likely to include many genes that are valuable as vaccine candidates and drug targets.

With the assumption that many proteins expressed at the infective larval stage facilitate transmission to the mammalian host, the aim of this thesis was to identify and characterise genes abundantly expressed at this stage.

CHAPTER 2

MATERIALS AND METHODS

Appendix 2 *E. coli* strains used in this study

Appendix 3 Primers used in this study

2.1 Parasite isolation

Vector stage parasites were obtained from *Aedes aegypti* mosquitoes infected with *B. malayi* by membrane feeding on human citrate-treated blood mixed with peritoneal-derived microfilariae to a concentration of 16,000 mf/ml. 12 – 14 days after infection infective larvae were recovered by gentle crushing of mosquitoes between a glass plate and a test tube, following which larvae were separated from mosquito debris in a Baermann apparatus containing Grace's insect tissue culture medium (Gibco-BRL, UK) maintained at insectory temperature (27°C).

Male jirds (*M. unguiculatus*) were infected intraperitonealy with infective larvae and used as a source of developing larvae, adult worms and microfilariae (Mf). Parasites were removed from the peritoneal cavity of jirds that had been euthanised by cardiac puncture under anesthesia, and washed in RPMI 1640 containing 25mM HEPES and glutamine (Gibco-BRL, UK). Mf were isolated from lavage fluid by gentle centrifugation (1000 rpm for 10 min) and purified by passage through PD10 Sephadex® G-25 M columns (Pharmacia Biotech AB, Uppsala, Sweden)

2. 2 Preparation of parasite antigen

B. malayi antigen was prepared by homogenization of mixed sex adult worms or Mf in PBS (Sigma, cat. no. P-4417, tablets giving 0.1M phosphate, pH 7.4) on ice followed by centrifugation at 12,000 x *g* for 20 min. The resultant supernatant was passed through a 0.2 μm filter prior to protein concentration determination by the Coomassie Plus protein assay (Pierce, Rockford, Illinois).

2. 3 Isolation of parasite RNA

Total RNA from all parasite stages was isolated using RNAzol™ or TRISOLV. Parasites were in a minimum volume of PBS were mixed with TRISOLV/RNAzol (1 volume sample: 9 volumes of TRISOLV/RNAzol) and

disrupted in an eppendorf homogenisers (Biomedix, Pinner, UK). To 0.5 ml of homogenate was added 100μl of chloroform (HPLC grade, 99.9%), shaken vigorously for 15 sec, placed at room temperature for 2–3 minutes, and then centrifuged at 12,000 g (4°C) for 15 min. The upper aqueous phase was transferred to the fresh tube, mixed with 250μl isopropanol (>99%, molecular biology grade Sigma), and stored at room temperature for 5-10 minutes. After centrifugation at 12,000 g (4°C) for 10 min, the RNA was washed with 1.5 ml of 75% ethanol by vortexing, and subsequently centrifuged for 5 min at 7,5000 g (4°C). At the end of the procedure, the RNA pellet was dried completely and immediately dissolved in 1 mM EDTA (pH 7) or autoclaved MilliQ water.

2. 4 Generation and cloning of abundant transcripts.

Total RNA was extracted from vector-derived infective larvae and larvae harvested from jirds 9 days and reverse transcribed using the GeneAmp RNA PCR kit (Perkin Elmer, CA, USA). Each reaction (20 µl) contained 2.5 µM DG d(T), 1mM of each dNTP, 1U of RNase inhibitor, 2.5U of MuLV reverse transcriptase. Reverse transcription was carried out at 42°C for 15 minutes with the poly $(A)^+$ tail complementary oligo-dT primer DG d(T) (Appendix 3) which contains 5' *Bam* HI and *Xba* I restriction sites. PCR was then carried out on the resultant first-strand cDNA using an SL1 oligo incorporating 5' *Not* I and *Xba* I restriction sites (Appendix 3) and *Pfu* DNA polymerase (Stratagene Inc. La Jolla, CA, USA). Thirty-five cycles of amplification were performed using a program of 94°C for 1 min 55°C for 3 min and 72°C for 15 min.

PCR products were identified by electrophoresis on a 3% Metaphor agarose gel (FMC, UK) followed by brief staining with ethidium bromide (0.5 µg/ml in water). Abundant bands were excised and purified using Geneclean (Bio 101, La Jolla, CA, USA). Purified bands were re-amplified and restriction digested with *Bam* HI and *Not* I together for 3 h at 37°C. Bands were then purified (Clean-up Kit, Promega, Wisconsin, USA). pBluescript KS (Stratagene, Appendix 4) was digested with *Bam*H I and *Not* I (10 U each enzyme for 4 h at 37°C), purified

(Clean-up Kit, Promega, Wisconsin, USA) and alkaline phosphatase-treated to reduce background from plasmid digested with only one restriction enzyme (1 unit for 15 min at room temperature, then 55°C for 15 min after adding another 1 unit of enzyme). The enzyme was removed (DNA Clean-up Kit, Promega) and ligation performed for 4 h at 16°C in the presence of 2 U of T4 DNA ligase. Ligated DNA was transformed into competent XL1-Blue *E. coli* (Stratagene) and recombinants identified by X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactoside)/IPTG (isopropyl-β-thiogalactopyranoside)/ampicillin selection.

^{32}P labelling of the RT-PCR products was carried out to allow better visualisation of individual bands. For this 4μl of RT-PCR product was re-amplified through 4 cycles in the presence of 40μCi ^{32}P-dCTP (Amersham International, Bucks, UK). After electrophoresis the gel was dried down overnight between layers of absorbent towels and autoradiographed.

2. 5 Southern Hybridisation.

To assess diversity of sequences amongst clones produced from each abundant band, inserts were PCR amplified with SL1 and DGDT primers. The PCR product of one clone was also labelled with digoxigenin (Boehringer Mannheim, UK) following the manufacturer's instructions and used to probe a Southern blot of all of the PCR products separated on a 2% agarose gel.

2. 6 RT-PCR from vector- and mammalian-stage parasites

RNA from infected mosquitoes was extracted using either TRIZOLV or RNAzol B (Biotex Inc). Total RNA from individual bloodfed mosquitoes was isolated without any attempt to recover the larvae. For each mosquito 40 μl of first strand cDNA was synthesized using the GeneAmp RNA PCR Kit (Perkin-Elmer) with the kits oligo $d(T)_{16}$ primer. To detect larvae within infected mosquitoes each first strand cDNA was amplified with primers specific for Bm-*cpi-2*. Total RNA from these positive mosquitoes was pooled and PCRs using gene-specific primers carried out using 5μl pooled vector stage cDNA or 1-2 μl mammalian stage first

strand cDNA. As a control, identical volumes of each cDNA was also amplified with primers specific for β-tubulin (Guenette et al., 1992). 25 μl PCR reactions were performed under standard conditions; (94°C for 1 min, 55°C for 1 min, 72°C for 1.5 min, 35 cycles; 72°C for 10 min) and included 50 μM of each primer, 0.2 mM each dNTP and 1.25 U Taq polymerase. Total RNA from larvae during the first 25 days of infection of the jird was reverse transcribed with the GeneAmp kit using the kits oligo $d(T)_{16}$ primer and amplified

2. 7 Isolation of parasite genomic DNA

Genomic DNA was isolated from *B. malayi* adult parasites. Parasites were homoginised in a pestle and mortar immersed in liquid nitrogen followed by digestion with proteinase K (Sigma) in lysis buffer (100 mM NaCl, 50 mM EDTA, 1% SDS, 1% 2-ME, 10 μg/ml RNase, 100 μg/ml proteinase K, 100 mM Tris-Cl, pH 8.5) at 65°C for 30 min with occasional inversion. Genomic DNA was extracted with 0.5 ml phenol by inversion for 5 min. After spinning 10 min at 12,000 rpm, the upper aqueous phase was extracted with 0.5 ml of phenol/chloroform and chloroform. After spinning 10 min at 12,000 rpm, 0.4 ml of the upper aqueous phase was added to 1ml of 100% ethanol (kept at -20°C). Precipitated DNA at the interface was hooked out and washed twice with 70% ethanol and once with 95% ethanol. The DNA pellet was briefly air dried and then re-suspended in TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0). Genomic DNA preparations were kept at 4°C.

2.8 PCR of genomic DNA

PCR was performed in 0.5 ml Eppendorf tubes. Twenty-five microliter reactions contained 1.25 –2.5 U Taq polymerase, 5μm each dNTP, 50ng each oligo. The amount of DNA and the cycling conditions differ for each template and primer combination. For most PCR reactions the following cycling condition were used: 94°C, 3 minutes, followed by 35 cycles of 94°C, 15 seconds, 55°C, 20 seconds, 72°C, 3 minutes, then 10 minutes at 72°C.

The products of the PCR reactions were mixed with DNA loading Buffer (6x stock solution: 30% glycerol, 0.25% bromophenol blue, 0.25% xylene blue) and analysed by electrophoresis in TAE (50x stock solution: 242g Tris base, 37.2 g $Na_2EDTA{\cdot}2\ H_2O$ or 100 ml of 0.5 M EDTA, 57.1 ml glacial acetic acid, d H_2O to 1 litre) on agarose gels with ethidium bromide staining (stock solution 50μg/ml).

2. 9 PCR of cDNA libraries

cDNA Libraries constructed from vector-derived L3, L4 (harvested 10 days after infection of jirds), adult male and microfilariae were provided by the Filarial Genome Project. PCR was carried out on 1 μl aliquots of each amplified library with gene specific primers listed in Appendix 3. Thirty-five cycles of amplification were performed using Promega Taq polymerase under the following cycling conditions: 94°C for 1 min, 55°C for 1 min and 72°C for 1.5 min.

2. 10 Cloning of PCR products

Taq polymerase amplified PCR products were ligated directly into pGEM-T (Promega, Appendix 4). Ligations were carried out following the manufacturer's instructions. For each ligation reaction, 2 µl of gel-purified PCR product were mixed with 1 µl of T-vector, 1 µl of 10x ligation buffer (0.5 M Tris pH 7.4, 0.1 M $MgCl_2$, 0.1 M dithiothreitol, 10 mM spermidine, 10 mM ATP, 1 mg/ml BSA), 1 µl of T4 DNA ligase (3 Weiss Units/µl) and 5 µl of dH_2O. Ligation mixes were incubated at 15°C overnight, and 5 µl taken for transformation of chemically competent *E.coli* (XL-1 blue MRF´ or JM109).

2.11 Preparation of competent *E.coli* cells

A single colony of *E.coli* was added into 100 ml of LB broth in a 1 liter flask and incubated with shaking for approximately 3-6 h until the cell density was 4 - $7x10^7$ viable cells/ml (OD_{600} of 0.4-0.6). The culture was transferred into ice-cold 50 ml tubes, chilled on ice for 10 min and then centrifuged at 4000 rpm for 10 min at 4°C. The pellet was then re-suspended in the original culture volume of an

ice-cold $CaCl_2$/glycerol solution (60 mM $CaCl_2$ and 15% glycerol, filter sterilised) incubated on ice for 30 min. The cells were then collected by centrifugation as above, re-suspended in 1/2 of the original culture volume of an ice-cold $CaCl_2$/glycerol solution, and incubated on ice for 10 min. The cells were then collected by centrifugation again, re-suspended in 1/30 of the original culture volume of an ice-cold $CaCl_2$/glycerol solution and stored on ice overnight. The following morning the cell suspension was dispensed as 100 µl aliquots into prechilled eppendorfs and stored at –70°C.

2.12 Transformation of *E.coli*

Competent cells (50-100 µl) were mixed with either 5µl of a ligation reaction or 10ng of cut plasmid DNA. The cells were then incubated on ice for 30 min, heat-shocked at 42°C for 40 seconds and incubated on ice for 2 minutes. To each transformation mixture, 100-200 µl of room temperature SOC medium (GIBCO-BRL) was added and the mixture was incubated at 37°C for 1 hr. This mix was then plated out on LB agarose plates containing the appropriate antibiotic and incubated at 37°C overnight. For PCR products inserted into the coding sequence of β-galactosidase blue/white screening was carried out. Transformation mixes were plated out onto LB-agarose/antibiotic plates spread with 50µl X-gal (50 mg/ml solution in dimethyl-formamide) and 15µl 100mM IPTG. Plates were incubated overnight at 37°C and white colonies picked. Cells containing plasmids with inserts were identified by PCR of colonies. Picked colonies were transferred to 15 µl LB and 1 µl of this was used in 20 µl PCR reactions with either vector-specific or gene-specific primer pairs. Cycling conditions were those used for amplification of genomic DNA.

2.13 Clone excision

For EST clones obtained as phage (lambda Uni-Zap, Stratagene) plasmids were excised as follows: XL-1-Blue MRF´cells were grown overnight in LB medium then diluted 1/100 and grown for 2 – 3 hours to an OD600 of 0.2 – 0.5. The cells were then spun (1500g) and re-suspended to an OD600 of 1.0 in 10mM

MgSO4. 200μl of cells were then added to 250μl phage stock and 1μl ExAssist helper phage (>1 x 106 pfu/μl, Stratagene) and incubated at 37°C for 15 minutes. 3 ml LB was then added and the mixture incubated for a further 2 hours with shaking. pBluescript plasmid particles were then obtained by heating the mixture to 70°C for 15 minutes and centrifuging at 4000g for 15 minutes. 100μl supernatant was then added to 200μl SOLR cells (Stratagene) grown from a 1/100 dilution overnight culture to an OD600 1.0. After incubating at 37°C for 15 minutes 10–50μl were plated onto a LB- 50μg/ml ampicillin plate and incubated overnight at 37°C. Single colonies were then picked and grown overnight in LB-ampicillin. Plasmid was obtained using the QIAprep Spin Miniprep Kit (Qiagen).

2.14 Plasmid DNA purification

Plasmid DNA was prepared by QIAprep Spin Miniprep Kit (QIAGEN Inc.), according to manufacturer's instructions. Single colony or frozen culture bacteria were inoculated into 10 ml of LB broth containing the appropriate antibiotic and grown overnight at 37°C with shaking. The cells were collected by centrifugation at 3,000-4,000 rpm for 10 min. Plasmid DNA was store at 4°C or –20°C for long term storage.

2.15 DNA sequencing

Automated sequencing of plasmid DNA was carried out using dye terminators with the PRISM™ cycle sequencing kit (Applied Biosystems Inc.) and the 377 automated sequencer (ABI). Protocols were carried out as described by the manufacturer (ABI PRISM dye Terminator protocol, P/N 402078). For each reaction, the reagents were aliquoted into a 0.5 ml PCR tube as follows: 4 μl Terminator Ready Reaction Mix, 2 μl dsDNA Template (0.2 mg/ml), 0.5 μl sequencing primer (20-30 μg /ml), 3.5 μl H2O with total volume 10 μl. Cycle sequencing was performed on a HYBAID DNA thermal cycler (OmniGene) with the following condition: 25 cycles (96°C for 30 sec, 50°C for 20 sec, 60°C for 4 min, respectively). After completion, reaction products were mixed with 1μl of 3 M sodium acetate (pH4.6 or 5.2) and 25 μl of cold 95% ethanol, and placed on ice for

10 min. After centrifugation at 13,000 rpm for 20 min, the precipitated pellets were washed again using 125 µl of cold 70% ethanol and finally resuspended in 3 µl loading buffer (5 parts of deionised formamide to 1 part of 50 mg/ml Blue dextran in 25 mM EDTA, pH8.0). Samples were heated to 90°C for 3 min and 1.5 µl of each sample was loaded. 6% Denaturing gels were run at 50 W for 7 h. Chromatograms were viewed and data was edited using the Seqed™ program (ABI) and using MacVector program version 6.0. (Oxford Molecular Group, UK)

2. 16 Cloning into expression vectors

Recombinant proteins were produced in pET-29-T and pET-22 expression vectors (Novagen, Madison, WI, USA). In this family of related plasmids, expression of target genes is controlled by the bacteriophage T7 promoter. Since *E. coli* RNA polymerase does not recognise the T7 promoter, transcription of the gene is absence in host cells used to establish the plasmid thus increasing the chance of maintaining a plasmid expressing toxic proteins. Expression is achieved in host cells with an introduced T7 polymerase gene under the control of the lac promoter. Thus expression is induced by IPTG.

pET-29-T is a T-tailed member of the pET family (Appendix 4). Taq polymerase PCR products were gel purified and ligated directly into the vector with 2-3 U T4 DNA ligase. The ligation mix was transformed into chemically competent Xl-1 Blue cell and plated on to LB agar containing 30µg/ml kanamycin. To ensure selection of plasmids with inserts in the correct orientation for expression colonies were PCR screened with T7 promoter primer in combination with a 3′ insert-specific primer.

Bm-CPI-1 and -2 were cloned into pET-22 (Appendix 3) using *Msc* I and *Xho* I restriction sites (CPI-1) and *Bam* HI and *Xho* I sites for CPI-2. Primers incorporating these restriction sites (Appendix 3) were used to amplify the CPI fragments from a L3 cDNA library (SAW94WL-BmL3). pET-22 plasmid DNA and the PCR products were digested with the appropriate enzymes followed by purification using microconcentrators (Microcon, Millipore). Purified cut plasmid and cut insert DNA were ligated overnight at 16°C with 3-5 U T4 DNA ligase.

Colonies were PCR screened with a combination of T7 promoter primer and 3′ insert-specific primer as described above.

2.17 Recombinant protein expression in *E.coli*

pET29-T vector and pET-22s were used to express polypeptides fused with hexa-Histadine-tag in *E. coli* BL21 (DE3) host.
Two microlitres of ligation reaction containing recombinant plasmids were transformed into XL-1 blue MRF′ cells then grown on LB-kan/amp agar plates overnight at 37°C. Positive clones were identified by direct PCR. Plasmid DNA was prepared with the QIAprep Spin Miniprep Kit (QIAGEN Inc.) and sequenced to provide confirmation of identity. Positive plasmid DNA with correct sequence was transformed into BL21 (DE3) cells, and grown on LB-kan/amp agar plates overnight at 37°C.

Single colonies were picked and put into 10 ml of LB-kan/amp and grown at 37°C with shaking to an OD_{600} of 0.6-0.8 (about 3-6 h). 5 ml of culture was removed as a control before induction, and to the remaining culture 1 M IPTG was added to a final concentration of 1mM in order to induce expression of the fusion protein. Cells were then grown for a further 3-6 h. Cells were collected by centrifugation at 5,500 rpm for 10 minutes at 4°C and resuspended in 1 ml of ice-cold 1x PBS. The cells were lysed by sonication on ice and the lysate microcentrifuged. A sample of both the supernatant and pellet was analysed by SDS-PAGE and Coomassie staining to indicate the expression and solubility of the His•tag fusion protein.

2.18 Purification of recombinant proteins

Preparing the cell extract. 1 litre of bacteria expression culture was harvested by centrifugation at 10,000 - 20,000 x g for 20 min, resuspended in 20 ml ice-cold Binding buffer (8X Binding buffer: 40 mM imidazole, 4 M NaCl, 160 mM Tris-HCl, pH 7.9) and sonicated with the sample on ice. The lysate was centrifuged

at 10,000 - 15,000 x g for 15 min and the supernatant was filtered through a 0.45 mm Millex-HV Filter Unit (Cat No SLHVR25LS, MILLIPORE) to prevent clogging of the resin

Preparing the resin. A few ml of deionized water was added to the dry column (PIERCE) and the column flowing was started. The bottle containing His-Bind resin (Novagen) was gently mixed by inversion and 5 ml of resin was then added into column, which can be used to purify up to 20 mg of target protein. When the level of storage buffer (20% ethanol) drops to the top of the column bed, the column was washed with the following sequence: 7.5 ml of deionized water, 12.5 ml of 1X Charge buffer (8x Charge buffer: 400 mM $NiSO_4$), 7.5 ml of 1X Binding buffer.

Performing column chromatography. When the Binding buffer drains to the top of the column bed, the column was loaded with the prepared extract, washed with the following sequence: 25 ml of 1X Binding buffer, 15 ml of 1X Wash buffer (8X Wash buffer: 480 mM imidazole, 4 M NaCl, 160 mM Tris-HCl, adjust final pH to 7.9), 25 ml of 1X TBS (150 mM NaCl, 20 mM Tris-HCl, pH 7.9). Recombinant proteins were eluted with 15 ml of TBSE (TBS containing 50 mM EDTA), injected into dialysis cassettes (PIERCE) and then dialysed in TBS for 24 h at 4°C.

2. 19 Detection of cysteine protease inhibition by recombinant CPIs

A fluorogenic assay was used to detect inhibition of papain and cathepsin B by both CPI's (Katunuma and Kominami, 1995). Assays were carried out at pH 5.5 in 0.3 M acetate buffer containing 0.5mM DTT and either 0.15μM papain (Calbiochem) or 0.1μM bovine cathepsin B (Sigma). Enzyme solutions were pre-incubated with 0-1.5 μg of purified recombinant protein in a total volume of 1ml at 37°C for 10 minutes after which 10μM Z-Phe-Arg-AMC substrate (Bachem, UK) was added. Increases in fluorescence were measured over 2 minutes using a Perkin Elmer LS50B spectrofluorimeter using excitation and emission wavelengths of 360 and 440 nm respectively. Initial reaction rates were calculated for each concentration of recombinant protein used as the change in fluorescent units/minute. As a negative control, *Bm*-ALT-1, an unrelated recombinant protein

B. malayi expressed with identical fusion partners, was tested under identical assay conditions.

2.20 Production of antisera in mice

6-8 week old mice were generally purchased from Harlan-UK (Bicester, UK), or obtained from source (Ann Walker House, Edinburgh University). For immunisation, mice were injected subcutaneously with 20μg recombinant protein in complete Freunds adjuvant and boosted 1 month later with 10μg antigen in Freunds incomplete adjuvant. Mice were bleed for antisera 1 month after the boost. For each antigen 6 BALB/c mice were immunised. To reduce the risk of genetic restriction in the response to the antigens, 6 CBA mice were also immunised with CPI-1, CPI-2 and ALT-1.

2.21 Vaccination

Male jirds (*M. unguiculatus*) were immunized subcutaneously with 75 μg of rALT-1 in Complete Freund's Adjuvant (CFA) or with CFA alone (6 jirds/group). At weeks 32 and 33, boosts were given of 25 μg rALT-1 in Incomplete Freund's Adjuvant or IFA alone, and at week 45 a final boost of 7.5 μg in IFA was given. Two weeks later, all jirds were challenged with 300 larvae of *B. malayi* introduced intraperitoneally. After 4 weeks, jirds were euthanized and parasites recovered from the peritoneal cavity and testes. Parasite recoveries and counting were performed without knowledge of the experimental status of each animal.

2.22 SDS-PAGE

Proteins were resolved on 15% SDS-PAGE gels (Acrylamide:bis-Acrylamide ratio, 74:1). The stacking gel consisted of 5% acrylamide/0.125M Tris/0.1%SDS and the main (separating) gel contained 15% acrylamide/0.375M Tris-HCl/0.1% SDS. Gels were electrophoresed at 20mA in running buffer containing 0.38M glycine/ 0.5M Tris/ 0.1% SDS.

Separated proteins were then either stained with Coomassie blue staining solution (50% methanol, 10% acetic acid, 0.25% Coomassie Blue R-250) or electrophoretically transferred onto a membrane for Western blotting.

2.23 Western blotting

Proteins separated by SDS-PAGE were transferred onto PVDF (Immobilon, Millipore) or Hybond-C nitrocellulose (Amersham) membrane. Proteins were transferred to the membrane in a LKB Novablot apparatus in transfer buffer (0.19M glycine, 0.025M Tris base, 0.05% SDS, 10% methanol). Transfer was conducted at 0.8mA/cm^2 gel. Transferred proteins were visualised by Ponceau S solution (cat. no. P-7170, Sigma).

The membranes were incubated in blocking solution (5% fat-free dried skimmed milk, MARVEL) in PBS, pH 7.4) overnight, rinsed once in PBS-TT (0.1% Triton X-100 and 0.05% Tween-20 in PBS), and then incubated with diluted antibody (1:100-500) in blocking solution containing 0.1% Triton X-100 and 0.05% Tween-20 for 1 h at room temperature. After washing with PBS-TT, the membranes were incubated with peroxidase-conjugated rabbit anti-mouse IgG (BioRad, UK). Bound antibodies were detected by chemiluminescence using ECL$^+$Plus Western blotting detection system (Cat No RPN2131, Amersham Life Science) and Hyperfilm ECL (Cat No RPN3103H, Amersham Life Science) following the manufactures instructions.

2.24 Metabolic labeling of parasites

Approximately 50 mixed sex worms or 500 larvae were maintained at 37°C in RPMI supplemented with 25mM HEPES, 1% glucose, 2mM L-glutamine and 100 U/ml penicillin, 100μg/ml streptomycin and 1mCi Trans^{35}S-label (ICN, CA, USA). During the culture period a solution of 7.5% sodium bicarbonate was added each day to maintain a neutral pH and glucose was added to a final concentration of 1% every two days. After 5 days of culture the medium was centrifuged at 12,000g to remove microfilariae and filtered through a 0.22μm filter.

2.25 Surface-specific labelling of parasites

Surface exposed molecules of adult and L2 parasites were labelled by the Iodogen reagent (Pierce, Rockford, IL, USA), which acts as surface-specific catalyst as it is insoluble in aqueous solutions. In the presence of sodium ^{125}Iodine Iodogen introduced radio-iodine onto tyrosine residues. One hundred microlitres Iodogen (1 mg/ml in methylene chloride) is added to a 1.5 ml eppendorf tube and allowed to dry completely. Parasites are transferred to the tube in PBS together with 500μCi sodium ^{125}Iodine (IMS 30, Amersham, UK) and incubated at room temperature for 10 minutes with occasional agitation. Ten microlitres of saturated tyrosine solution is then added to quench the reaction. The parasite are removed from the tube and washed extensively in PBS.

2.26 Immunoprecipitation

Immunoprecipitations were carried out on 500μl of culture supernatant pre-cleared by incubation with 50μl formalin-fixed *Staphylococcus aureus* (Panasorb, Calbiochem, Nottingham, UK) for 15 minutes. After removal of Panasorb by centrifugation at 12,000g the cleared supernatant was incubated with 20μl antisera overnight at 4°C. Immune complexes were absorbed onto Protein G-Sepharose (Pharmacia) by rotating for 2 hours at 4 °C. The complexes were then washed in phosphate buffered saline containing 0.5% Triton X-100 and released by boiling in SDS-PAGE sample buffer. Proteins were analyzed by SDS-PAGE on 17.5% polyacrylamide gels.

2.27 Phylogenetic Analysis

Cystatin sequences, omitting their signal peptides, were aligned with the aid of the ClustalW function of MacVector with final manual adjustment. Using two stefin sequences as the outgroup, phylogenetic analyses using maximum parsimony were performed using PAUP 3.1.1, written by David L. Swofford. Bootstrap

support for the shortest tree found by heuristic searching was assessed using 100 replicates.

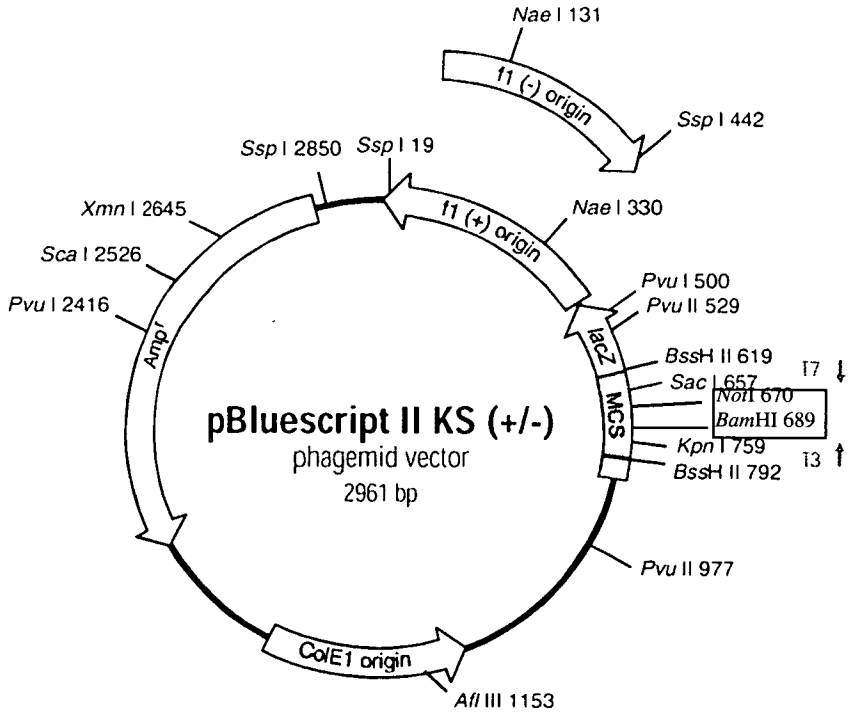
Appendix 1

Plasmids used in this study

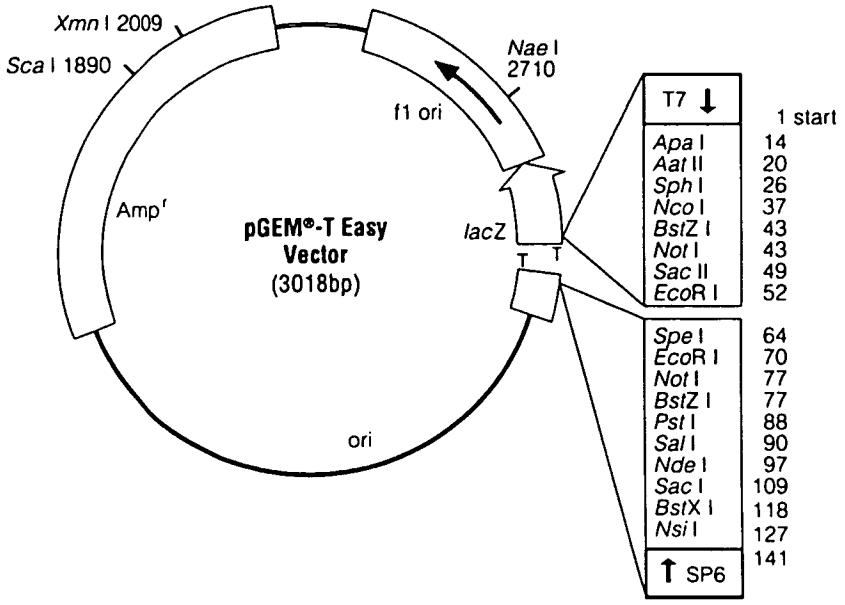
Name (company)	Brief description
pBluescript II KS(Stratagene)	Amp. Resistant. Blue/white screen
pGEM-T(Promega)	cloning of PCR products Amp. Resistant. Blue/white screen
pET-29-T(Novagen)	direct cloning and high level expression of PCR products. Fusions contain N-terminal S•Tag and C-terminal His•Tag. Kan. Resistant.
pET-22(Novagen)	High level expression of proteins fused to the pelB signal sequence for potential periplasmic localisation. C-terminal His•Tag. Amp. Resistant.

Vector map of plasmids used in this study

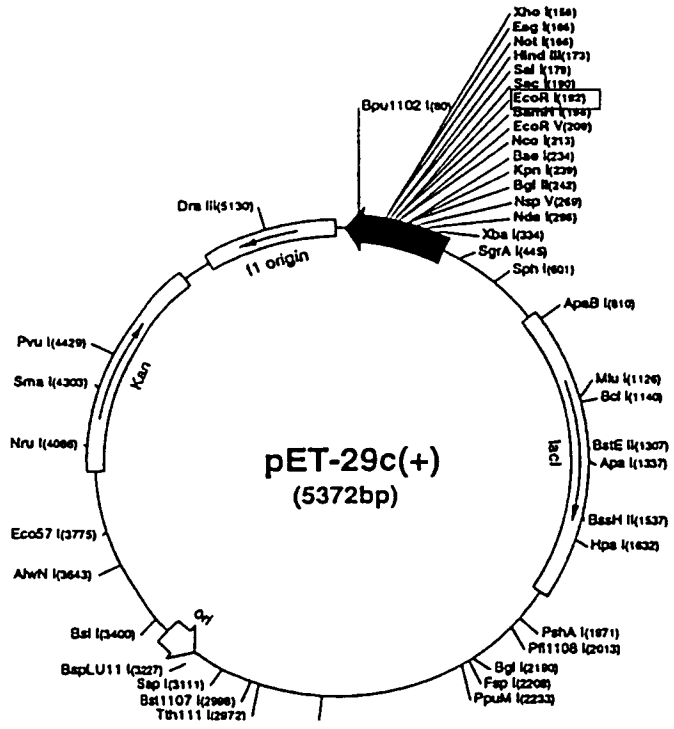
A



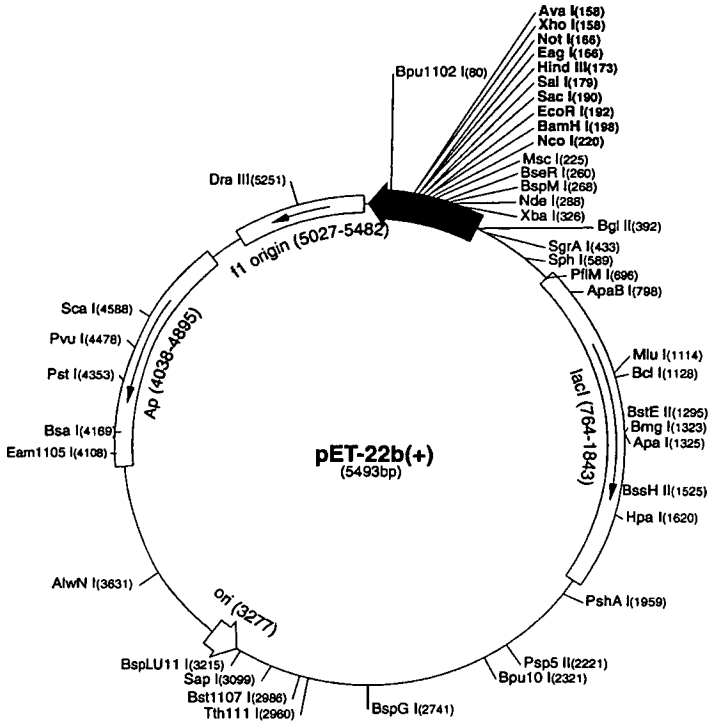
B


Xmn I 2009
Sca I 1890
Nae I
2710
f1 ori
Amp^r
pGEM®-T Easy
Vector
(3018bp)
lacZ
T T
ori
T7
1 start
Apa I 14
Aat II 20
Sph I 26
Nco I 37
BstZ I 43
Not I 43
Sac II 49
EcoR I 52
Spe I 64
EcoR I 70
Not I 77
BstZ I 77
Pst I 88
Sal I 90
Nde I 97
Sac I 109
BstX I 118
Nsi I 127
141
SP6

C


pET-29c(+)
(5372bp)
f1 origin
Kan
lacI
ori

D



A. pBluescript II KS(+). *Not I* and *Bam HI* sites were used to clone abundant transcripts. **B**. pGEM-T. A commercially derivative of pGEM 5Z f(+) prepared by restriction digesting with *Eco*R V and adding 3´ -T overhangs to allow the direct cloning of PCR products produced by Taq polymerases. **C**. pET-29-T. A T-ended expression vector also produced by digesting with *EcoR* V and T-tailing. Allows the production of recombinant proteins fused to a C-terminal S•Tag and N-terminal hexahistidine tag. **D**. pET-22. Allows the production of recombinant proteins fused to an N-terminal pelB signal sequence for potential periplasmic localisation plus a C-terminal hexahistidine tag. Inserts are introduced into both pET plasmids in a region transcribed by T7 RNA polymerase that is under the control of the host cell *lac* promoter.

Appendix 2

E. coli *strains used in this study.*

Strain(company)	Genotype	Brief description
XL-1 Blue(Stratagene)	*Δ(mcrA)183 Δ(mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 Lac [F'proAB lac*Iq*lacZ*ΔM*15*Tn*10(tet*^{r}*)]*	Tet resistant
JM109(Promega)	*end*A1, *rec*A1, *gyr*A96, *thi*, *hsd*R17 (r_k^-, m_k^+), *rel*A1,*sup*E44, Δ(*lac-pro*AB), [F´ *tra*D36, *pro*AB, *laqI*ZΔM15]	No antibiotic resistance
BL21(DE3)(Novagen)	*F- ompT hsds$_B$(r_B-m_B-) gal dcm (DE3)*	High level expression from plasmids containing the bacteriophage T7 promoter

Appendix 3 – Primers used in this study. Sequences underlined indicate restriction sites used for cloning. Extra nucleotides used for cloning into pET-29-T are bolded

Primer Pairs for expression cloning

Name (restriction site)		Location
4.1 pET For	**A**GATGACGAATTCGACGACGAATCCTCA	100-127
4.1 pET Rev	**A**CGAGCATTGCCAACTTTCTTGAGGCGC	380-403
061 pET For	**AA**TACGATTAAGCAGAGCTACATATAAT	86-112
061 pET Rev	**T**TGTCAAAGTGACTTTGAATGTGTTTTG	377-403
CPI-1 pET22For (*Msc* I)	GGCGATGG▼CCATGAGGGAAATACGATTAAGCAGAGCT	80-103
CPI-1 pET REV (*Xho* I)	CAC▼CTCGAGTGTCAAAGTGACTTTGAATGTGTTTTG	377-403
038 pET For	**T**TGATTCATCGACGAGAAATTCCTCAT	131-157
038 pET Rev	**G**AGAGTACCTTTGTCTCCAAAATATTC	505-531
CPI-2 pET22For (*Bam* HI)	C▼GGGATCCCTTGATTCATCGACGAGAAATTCCTCAT	131-157
CPI-2 pET22Rev (*Xho* I)	CAC▼CTCGAGTACTGACGAGTACCTTTGTCTCCAA	512-535
A033 pET For	**A**TTTGGTTGGGGATACTCATATTACGGT	107-133
A033 pET Rev	**T**TTTCTTCTTGCCCAGTACAGCACCAGCGAT	185-214
TPH pET For	**AA**TGTTGATCTTCAAGGATGCATTCACA	63-89
TCTP pET Rev	**T**TTGTTTTTCTTCAATGAGTGCCTCCTT	579-603

Primer Pairs for RT-PCR and library PCR

		Location
4.1 3′UTR (Used with 4.1 pET For)	TTGTTTTGCTTGCTTTGTAAGCATTTA	407-432
alt-2 For	GACGAAGAGTTCGATGACTCCGCAGCC	102-128
alt-2 3′ UTR	GTAGTATCAAAGACTGATTCATTCCTA	417-443
061 3′	CACAATCTGTTATTTATGAATTATCG	461-486
061 5′	ATGTTCTTCCCGATTGTATGG	23-43
038-5′	ACCAACTGTGTAAACAAAATAAGC	23-46
038-3′	ACCAAATCCAAATATTTATGCACCA	670-693
A033 3′ UTR (Used with A033 Pet For)	CGTATGACCTTTTCAACACTTAAAAAT	282-308
Tubulin A	AATATGTGCCACGAGCAGTC	121-140
Tubulin B	CGGATACTCCTCACGAATTT	409-428
Ce-*alt* For	GGAATGCCTTGCGAAACAGACCAAGAT	1-27
Ce-*alt* Rev	ATGTTTTGGTGGTGAGCAATGGTCCAA	211-237

Primers for genomic sequencing

		Location
g061-1 For	TGCACCGACCACCTGTAGAAA	362-382
g061-2 Rev	GTCGGTTTACTGCTGTCGATA	687-707
g038-1 Rev	CTGCAAAAAATCTTCCCATGG	1469-1489
g038-1 For	GGACAAATGCAGCGTGGCCAA	148-168
g038-E2 For	ATGCCAATCAAATTGCTAAAAGTTTCA	588-614
g038-3′	ACCAAATCCAAATATTTATGCACC	1657-1679
alt-2g Rev	GTTCTTTCGATGACGCACGAATGCAAC	990-964
alt-2E2 For	GCTCTTCGCACGAAGCTTGCTACGATC	530-556

Miscellaneous primers

		(Restriction enzyme site)
SL-1	GCTCTAGAGCGGCCGCGGTTTAGTTACCCAAGTTTGAG	(*Not* 1)
DG d(T)	AATTCGGATCCCCCGG(T)$_{18}$	(*Bam* HI)
T7	ACTATAGGGAGCTAAGCTTGG	
T3	ATTAACCCTCACTAAAG	
M13L	CGCCAGGGTTTTCCCAGTCACGAC	
M13R	AGCGGATAACAATTTCACACAGGA	

CHAPTER 3

Cloning of abundant *trans*-spliced cDNAs from L3 and L4 larvae

3.1 Introduction.

3.2 Results

- 3.2.1 Abundant PCR Products from vL3 and L4 cDNA
- 3.2.2 550 bp vL3 Products: Bm-*alt*-1, *cpi*-1, *rbp*-1 and a ribosomal protein.
- 3.2.3 500 bp vL3 Products: *glt*-1 and ribosomal proteins
- 3.2.4 300 bp vL3 and L4 products : *Bm-col*-3 and *slt*-1
- 3.2.5 750 bp vL3 band : *Bm-tph*-1
- 3.2.6 250 bp vL3 band : *Bm-efa*-1
- 3.2.7 150 bp vL3 band : *Bm-rps*-19
- 3.2.8 L4 abundant bands : 525 bp is *Bm-cdd*-1
- 3.2.9 Stage-Specific Expression of Abundant Transcripts

3.3 Discussion

INTRODUCTION

Experimental studies have shown that the transmission of filarial third-stage infective larvae (L3) from mosquitoes to their mammalian host and their subsequent development are promising targets for control of filarial infections. Entry into the mammalian host initiates recovery from an arrested state entered during the final stages of development in the vector. After transfer the larvae must quickly adapt to their new environment and prepare for the fourth moult 8-9 days later (Schacher, 1962). Besides moulting, the vertebrate-derived larvae undergo several changes in the definitive host to adapt to the new environment and evade immune responses, processes that require induction or down regulation of genes essential for their survival and development. Attempts have been made to define stage-specific proteins by comparing the total protein of polypeptides expressed by different life cycle stages. For example, this approach has been used to compare protein expression in filarial larvae as they re-initiate their development after transmission to the vertebrate host. These studies have concentrated on both surface (Apfel et al., 1992; Lal and Ottesen, 1988; Scott et al., 1990; Storey and Philipp, 1992)and secreted molecules (Bianco et al., 1990; Bianco et al., 1995; Frank and Grieve, 1991; Hong et al., 1993; Pogonoka et al., 1999; Richer et al., 1992). Information on these and other larval products will not only add to our understanding of the biology of infective larvae but may also identify potential targets for drug or vaccine-induced control of infection.

Despite this, few proteins have been cloned and characterised from the infective larval stage. In order to identify abundant, stage-specific transcripts in infective larvae, reverse transcriptase-polymerase chain reaction (RT-PCR) has been used to amplify total larval RNA using the conserved nematode spliced leader (SL1) and oligo(d)T. This technique has been used recently to identify abundant transcripts in *B. malayi* (Scott et al., 1995; Yenbutr and Scott, 1995) and *Toxocara canis* (Gems et al., 1994; Gems and Maizels, 1996). These studies have utilised the presence of SL1, a conserved 22 nt sequence *trans*-spliced onto the 5' end of many

nematode mRNAs, to amplify full length cDNAs in combination with a 3' oligo-dT primer (Blaxter and Liu, 1996; Nilsen, 1993). In both species several abundant bands were apparent when total SL/dT RT-PCR products were resolved on an agarose gel. With the assumption that highly expressed genes may play important roles in the development of larvae and their establishment in the mammalian host this technique has been used to identify abundant transcripts produced by infective third stage larvae harvested from the vector mosquito (vL3) 12 days after mosquito infection and fourth-stage larvae (L4) recovered 9 days after intraperitoneal infection of gerbils, immediately after the third moult. This chapter describes the cloning and identification by sequence similarity of each of the most abundant transcripts.

3.2.1 Abundant PCR Products from vL3 and L4 cDNA

Amplification of vL3 first-strand cDNA with SL and DGDT primers produced six distinct bands of 750, 550, 500, 300, 250 and 150 bp super-imposed on a smear of less abundant cDNAs on agarose gel electrophoresis (Figure 3.1). Parallel PCR of L4 cDNA produced only three visible bands of 300, 525 and 1.1 kb. Abundant bands were excised from the gel, amplified with a second round of PCR with the same primers and directionally cloned into pBluescript using *Bam H* I and *Not* I restriction sites present on the primers. The 1.1 kb band from day 9 larvae failed to re-amplify and thus was not analysed further. Multiple clones derived from each band were sequenced. A summary of the sequence analysis of clones produced by each abundant band is presented in Table 3.1.

3.2.2 550 bp vL3 Products : Bm-*alt*-1, *cpi*-1, *rbp*-1 and a ribosomal protein.

Seventeen clones derived from the intense 550 bp vL3 band were analysed and found to contain 4 distinct transcripts. Twelve clones represented one transcript, have been named *Bm-alt-1* (for abundant larval transcript-1). A detailed analysis of this gene and its protein product is presented in Chapter 5.

Three additional clones derived from the 550 bp band were found to encode a homologue of the cystatin family of cysteine proteinase inhibitors, which has been named *Bm-cpi-1*. Data on *Bm-cpi*-1 is presented in Chapter 4.

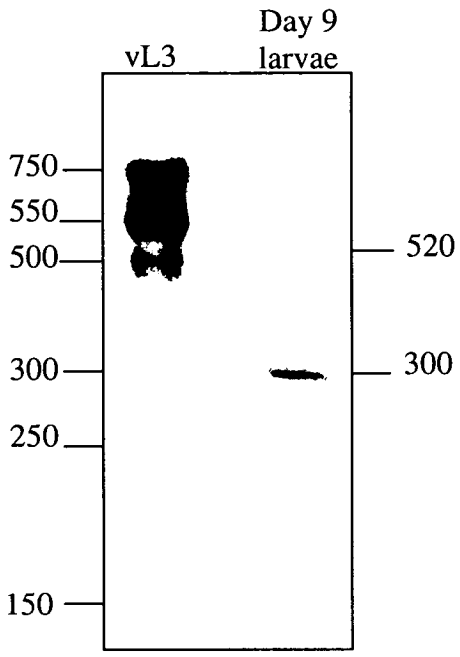


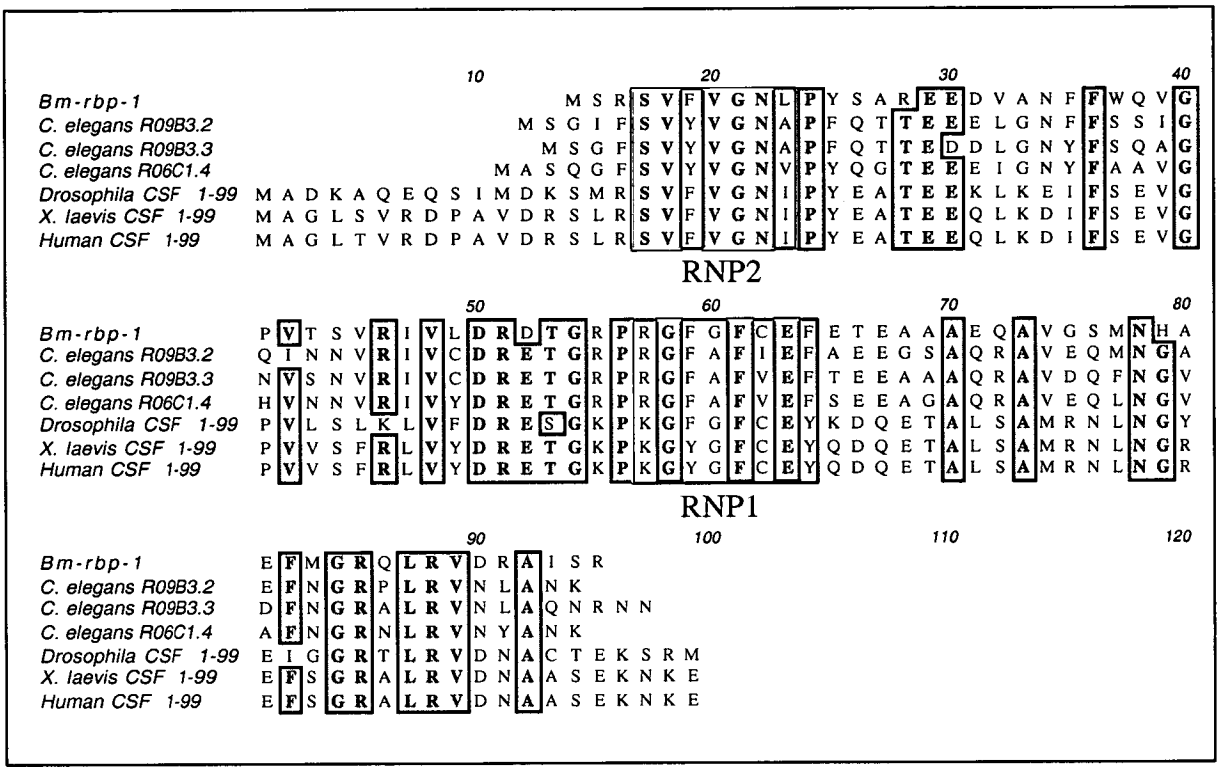
Figure 3.1: *Autoradiograph of total SL/oligo-dT RT-PCR products from mosquito-derived larvae (vL3) and larvae nine days after infection of jirds. Four microlitres of RT-PCR products were re-amplified using identical primer and cycling conditions in the presence of 40μCi [^{32}P]dATP.*

Size of band	*Name*	*Genbank accession*	*Identification*	*Relative Abundance**	*EST abundance‡* L3	L4
Day 0 L3 Bands						
750bp	*Bm-tph-1*	U80971	Translationally controlled tumour protein/Histamine releasing factor	8/8	0.20 (6)	0.99 (16)
550bp	*Bm-alt-1*	U57547	*D. immitis* larval 20/22 kDa protein homologue	12/17	1.39 (33)	0
	Bm-cpi-1	U80972	Cystatin-type cysteine proteinase inhibitor	3/17	0.38 (8)	0
	Bm-rbp-1	U80973	Polyadenylation factor/RNA binding protein	1/17	0.39 (12)	1.79 (29)
	Bm-rps-12	U81008	40S ribosomal protein S12	1/17	0.07 (2)	0.93 (15)
500bp	*Bm-glt-1* ¥	U80974	6 kDa Gly/Tyr-rich protein	8/11	0.7 (13)	0
	Bm-rpp-2	U80977	60S ribosomal protein P2	1/11	0.05 (1)	0.86 (14)
	Bm-rpl-44	U80976	60S ribosomal protein L44	1/11	0.05 (1)	0.74 (12)
			alt-1(truncated)	1/11		
300bp	*Bm-col-3*	U80975	Collagen	11/12	0.49 (15)	0.19(3)
	Bm-slt-1	U80978	Novel	1/12	0.29 (9)	0.06 (1)
250bp	*Bm-efa-1*	U84736	Elongation factor 1-α	15/15	0.26 (8)	0
150bp	*Bm-rps-19*	U80979	40S ribosomal protein S19	ND	0.13 (4)	0.25(4)
Day 9 L4 Bands						
525bp	*Bm-cdd-1*	U80980	Cytidine deaminase	8/8	0	5.9 (96)
300bp	*Bm-col-3*	U80975	Collagen	7/7	0.49 (15)	0.19(3)

Table 3.1. Summary of clones obtained from abundant bands. ** Relative abundance refers to the number of clones identified by sequence analysis or hybridisation compared to the total number of clones analysed. ‡ Number of matches with dbEST as a percentage of 3053 L3 and 1619 L4 EST's deposited into Genbank by the Filarial Genome Project as of 5.1.00. Total numbers of identical sequences are given in brackets. ¥ Originally termed Bm-alt-3, this sequence have been renamed glt-1 for glycine-tyrosine rich protein.*

Gene abbreviations are as follows:

tph, tumour protein homologue
alt, abundant larval transcript
cpi, cysteine proteinase inhibitor
rbp, RNA-binding protein
rps-12, ribosomal protein S12
rpp-2, ribosomal protein P2
rpl-44, ribosomal protein L44
col, collagen
slt, small larval transcript
efa, elongation factor-1α
rps-19, ribosomal protein S19
cdd, cytidine deaminase
glt, glycine/tyrosine rich



B

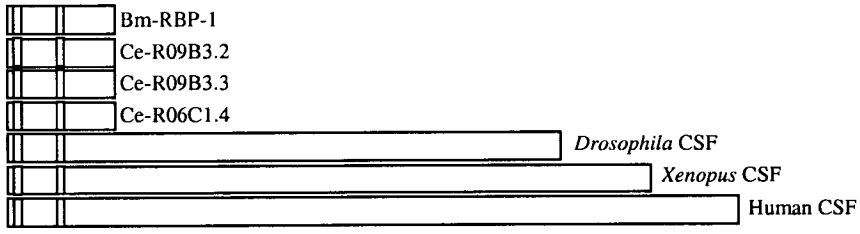


Figure 3.2: Panel A: Alignment of B. malayi RNA binding protein, rbp-1 with members of the cleavage stimulation factor/polyadenylation factor families. *GenBank accession numbers are as follows: Bm-rbp-1, U80973; C. elegans cosmid R09B3, Z81108; C. elegans cosmid R0C1, Z81106; Drosophilia CSF, AF170082; Xenopus leavis CSF, U17394; Human CSF, Z78021. The RNA recognition motif (RNP) is boxed in red. Note that only the N-terminal 99 residues of the Drosophilia, Xenopus and human sequences are shown in the alignment.* ***Panel B*** *shows that members of the cleavage stimulation factor/polyadenylation factor families are considerably longer and contain auxiliary motifs.*

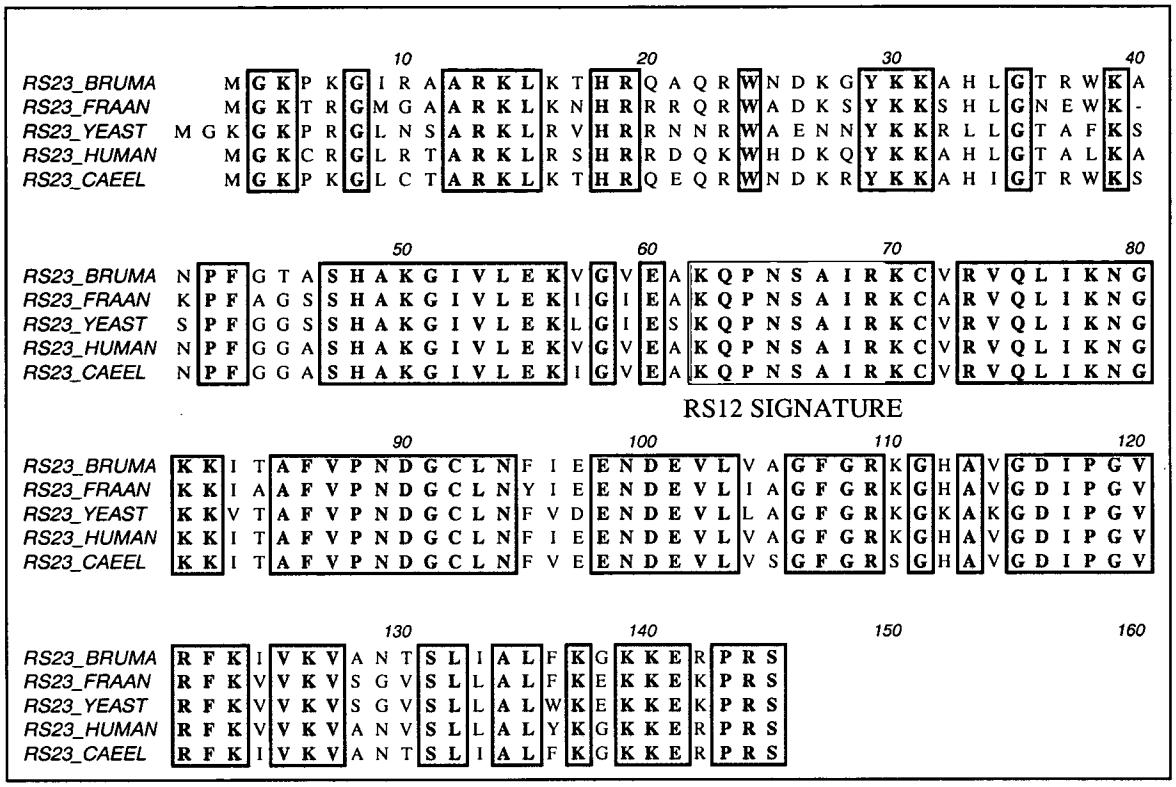


Figure 3.3: Alignment of Bm-RPS-12 (RS23_BRUMA) with members of the S12/23 family of ribosomal proteins. *Accession numbers are as follows: RS23_BRUMA, Swiss-Prot P907076; RS23_FRAAN, Swiss-Prot P46297; RS23_YEAST, Swiss-Prot P32827; RS23_HUMAN, Swiss-Prot P39028; RS23_CAEEL, Swiss-Prot Q19877. The RS12 signature is boxed in red.*

Of the two remaining 550 bp clones analysed, one is homologous to a partial cDNA recently identified in *B. pahangi*, and encodes a predicted protein with extensive similarity to a large family of RNA-binding proteins containing the ribonucleoprotein 1 and 2 consensus sequences (RNP-1 and RNP-2) Which appear to be directly responsible for the interaction of the protein with RNA. This transcript is therefore termed *Bm-rbp-1* (for RNA-binding protein-1). The predicted protein shares the greatest similarity to the RNA binding segment of cleavage stimulation factors (CSF)/polyadenylation factors (Figure 3.2A). These are proteins of around 64kDa and are among the factors required for polyadenylation and 3´ cleavage of pre-RNAs. In addition to CSFs these RNPs are found in more than 1000 proteins catalogued by SwissProt and consist of single or multiple N-terminal RNA recognition motifs in combination with other domains thought to direct other interactions (Takagaki et al., 1992). Other CSF domains include regions rich in one or two amino acids (predominantly Gly and Glu or Arg/Ser, Gly/Pro,). Additionally, a feature of human CSF is a repeat sequence of 5 amino acids (MEAR$^{A}/_{G}$) possibly mediating protein-protein interactions. CSFs are typically 400-600 residues in length, considerably longer than *Bm*-RBP-1 which has a calculated molecular weight of 9.2 kDa and contains a single RNA recognition motif with few additional residues, which are unlikely to contain auxiliary motifs (Figure 3.2B). Of the homologues found in the database three predicted genes from *C. elegans* are small and contain only the RNP motifs as the *B. malayi* gene. (Figure 3.2B). *Bm-rpb*-1 encodes an in frame stop codon at bases 322 - 324. The codon is conserved amongst the ESTs representing this gene. The EST sequences do not extend into the poly A tail, it is therefore difficult to assess whether *Bm*-RBP-1 represents the full length sequence, however the presence of the three *C. elegans* genes also containing a single RNA binding motif would suggest that this represents the full length of a novel group of nematode proteins. Although the function of these nematode genes is not yet known, their similarity to eukaryotic and plant proteins suggests that they play a role in RNA processing and gene expression (Heintzen et al., 1994; Nishiyama et al., 1997; Sturm, 1992; Takagaki et al., 1992).

The second 550 bp sequence is homologous to genes encoding the 40S ribosomal proteins S12/23 and is designated *Bm-rps*-12. One characteristic of this group of proteins from the small ribosomal is their high content of basic amino acids. *Bm-rps*-12 has a calculated PI of 11.0 and contains the ribosomal protein S12 signature: [RK]-x-P-N-S-[AR]-x-R (Figure 3.3).

3.2.3 500 bp vL3 Products: *glt*-1 and 60S ribosomal proteins P2 and L44

Of 11 clones obtained from the vL3 500 bp abundant band, 8 represented a novel transcript with an unusually long 3' untranslated region, which has been termed *Bm-glt-1*. This gene is further characterised in Chapter 6.

The remaining clones analysed from the 500 bp vL3 band encoded ribosomal proteins L44 (*Bm-rpl-44*, Fig. 3.4) and P2 (*Bm-rpp-2*, Fig. 3.5), and a truncated form of *alt-1* terminating at an A-rich stretch of sequence in which 11/18 bases are As. It appears that mis-priming of the oligo-dT primer produced this clone.

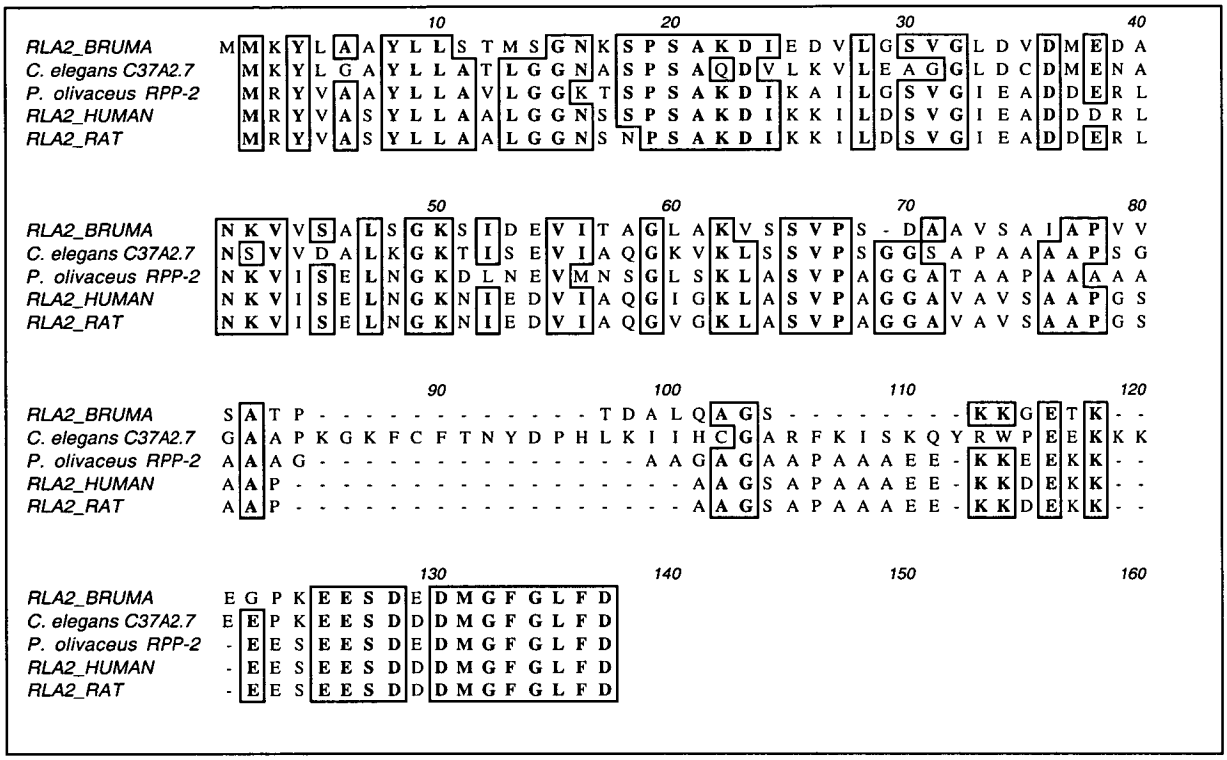


Figure 3.4 Alignment of Bm-RPP-2 (RLA2_BRUMA) with members of the P2 family of ribosomal proteins. *Accession numbers are as follows: RLA2_BRUMA, Swiss-Prot P90703 ;C. elegans C37A2.7, TrEMBL O01504; Paralichthys olivaceus RPP-2; RLA2_HUMAN, Swiss-Prot P05387; RLA2_RAT, Swiss-Prot P02401.*

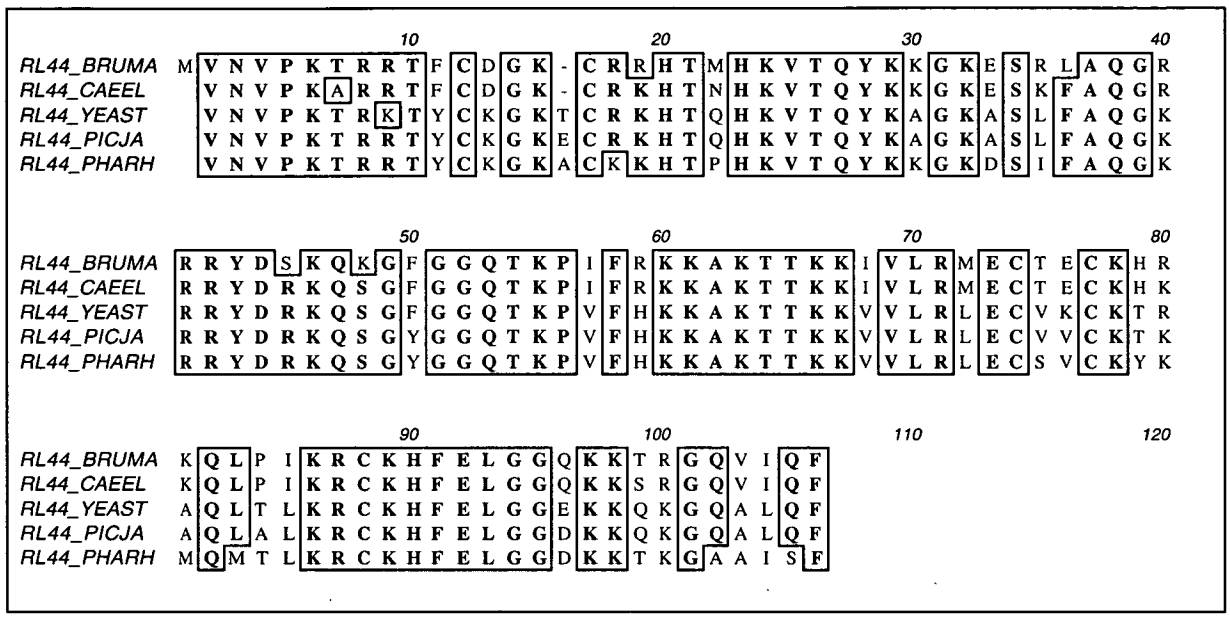


Figure 3.5: Alignment of Bm-RPP-44 (RL44_BRUMA) with members of the L44 family of ribosomal proteins. *Swiss-Prot accession numbers are as follows: RL44_BRUMA; P90702, RL44_CAEEL; P48166, RL44_YEAST; P02405, RL44_PICJA; P52809, RL44_PHARHA; O59870.*

3.2.4 300 bp vL3 and L4 products : *Bm-col*-3 and *slt*-1

Clones derived from the 300 bp bands from both vL3 and L4 represent a transcript with similarity to the signal peptide and part of domain I of the *C. elegans* collagens COL-12 and COL-13 (Johnstone, 1994). Based on this similarity this transcript is designated *Bm-col-3* (Fig. 3.6). *Bm-col-1* and *col-2*, which represent distinct collagen genes, have recently been described. The small size of this cDNA compared to other members of the family, the absence of an in-frame termination codon and the absence of (Gly-X-Y)$_n$ motifs, typical of collagens, suggest that the clone may have been truncated by mis-priming of the oligo-dT primer. Seven ESTs identical to *Bm-col*-3 can be found in dbEST, a consensus of which extends the sequence by 81 nucleotides. This extended sequence reveals that the SL/dT amplified 300 bp clone is truncated at an A-rich stretch (14/18 bases are A), and that the overlap region is identical at the amino acid level to the recently described *B. pahangi* collagen gene *Bp-col-1* (Bisoffi and Betschart, 1996). *Bp-col-*

1 appears to be 5´ truncated as the start of its translated sequence aligns with residue 72 of the *Bm-col*-3 protein

A single clone was also isolated from the L3 300 bp band for which no identity can yet be given. The longest open reading frame in this sequence is only 26 amino acids and the sequence is therefore designated *Bm-slt-1* (for small larval transcript-1) (Fig. 3.7). No database similarities were found other than with a family of highly similar ESTs from *Brugia* L3s. The nucleotide composition is heavily A/T biased - 74.5% of nucleotides are A or T, a percentage more typical of introns in *C. elegans* (Blumenthal and Steward, 1997) and *Brugia* (Hammond, 1994). The EST sequences support the sequence of *Bm-slt*-1 as many show the presence of the SL sequence at the 5´ end and also confirm the site of addition of

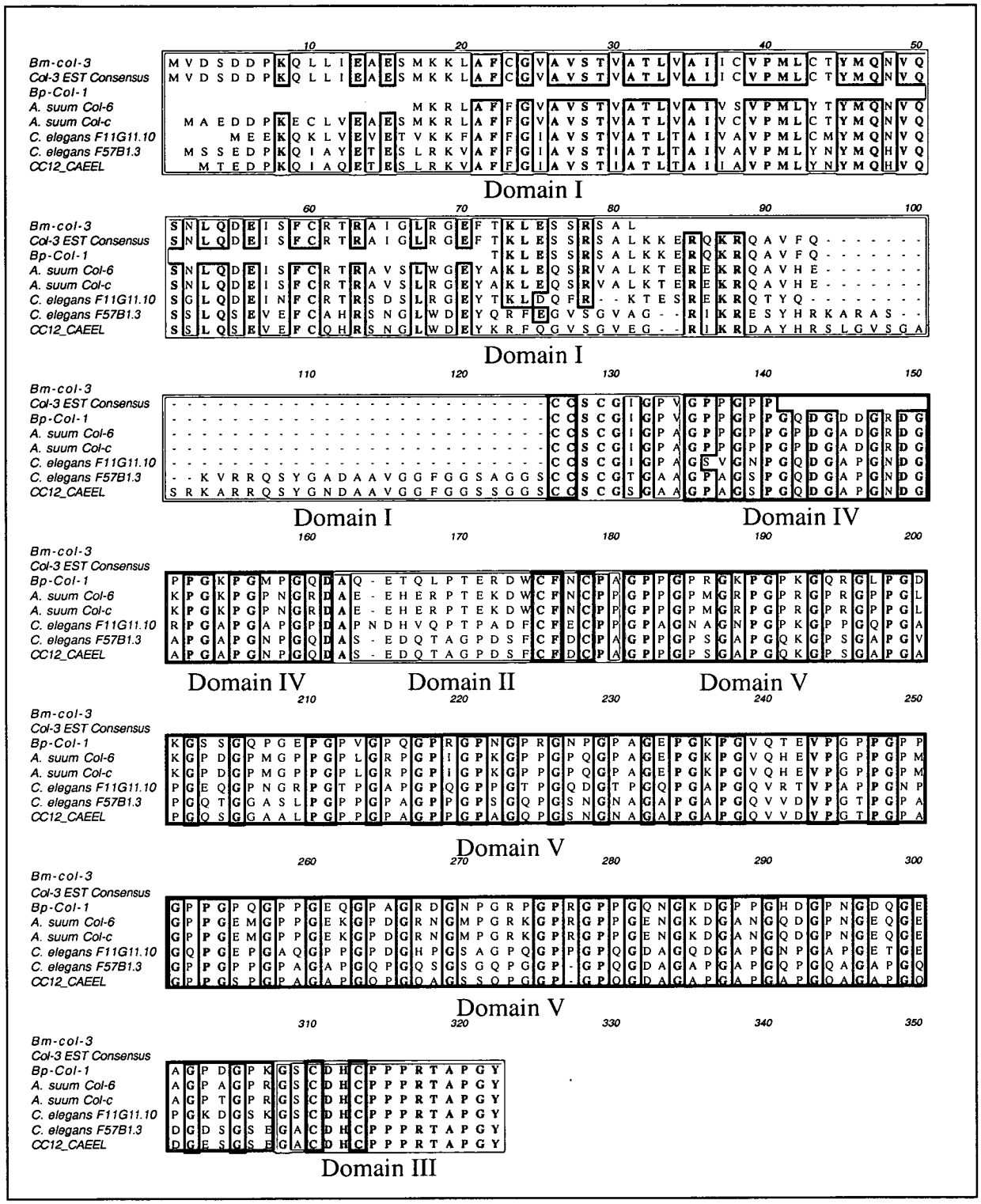


Figure 3.6. Alignment of Bm-COL-3 with its closest collagen homologues. *GenBank accession numbers are as follows: Bm-COL-3, U80957; B. pahangi, X92099; A. suum Col-6, AF035410; A. suum col C, AJ243805; C. elegans cosmid F11G11, U80451; C. elegans cosmid F57B1, Z78064; CC12_CAEEL, Swiss-Prot P20630. Collagen domains are boxed: Gly-X-Y domains in blue, others in red. Conserved cysteines are boxed in green. The col-3 consensus sequence was generated from 7 B. malayi ESTs (accession numbers) AI574499, AW159925, AW179743, AW225450, AW257696, AW043521 and AW 179891).*

the poly A tail. Taken together these data suggest that *slt*-1 is a real gene and not an artifact of the PCR amplification, although identification of the native protein will be required to provide conclusive evidence of this.

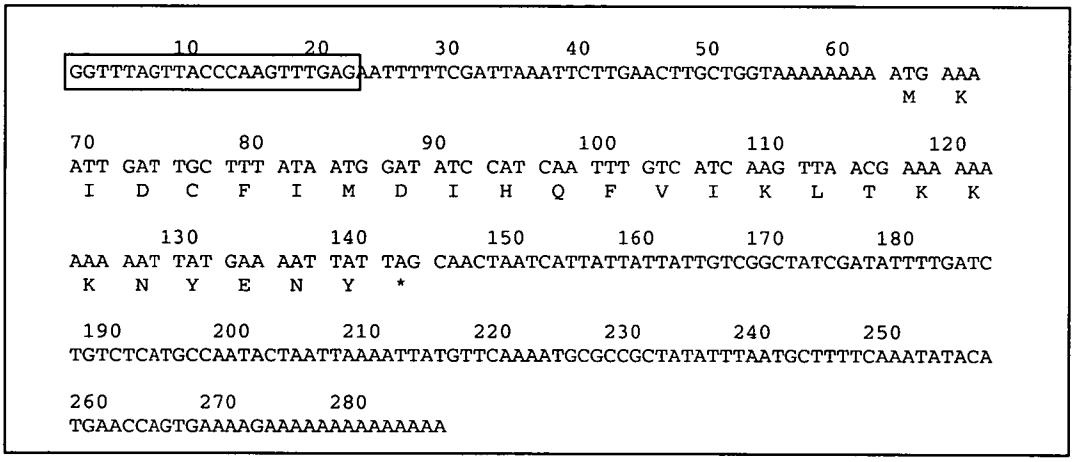


Figure 3.7: Nucleotide and deduced amino acid sequence of Bm-slt-1. *The nematode trans-spliced leader sequence is boxed.*

3.2.5 750 bp vL3 band : *Bm-tph*-1

Clones produced from the vL3 750 bp band all appear to derive from the same gene. Database searches revealed extensive similarity to translationally controlled tumour protein (TCTP), the sequence of which has recently been identified as an IgE-dependent histamine releasing factor (MacDonald et al., 1995). Characterisation of this gene and its protein product, termed *Bm-tph*-1 for tumour protein homologue), are presented in Chapter 7.

3.2.6 250 bp vL3 band : *Bm-efa*-1

The majority of clones derived from the vL3 250 bp band encoded a protein with similarity to elongation factor-1α (EF-1α) from a wide range of organisms. Comparison of this sequence with its homologue from *O. volvulus* (Alarcon and Donelson, 1991) shows that this is a truncated product containing only the 5′ end. (Fig. 3.2.8). EF-1α catalyses the GTP-dependent binding of aminoacyl-transfer

RNA to ribosomes and in eukaryotes it is the second most abundant protein after actin. It is the subject of recent interest because although EF-1α mRNA levels correlate with levels of cellular proliferation, it is found in vast molar excess to the other essential components of the translation machinery. This has prompted the suggestion that EF-1α may serve other purposes within the cell such as interaction with the cytoskeleton via its binding to actin (Condeelis, 1995).

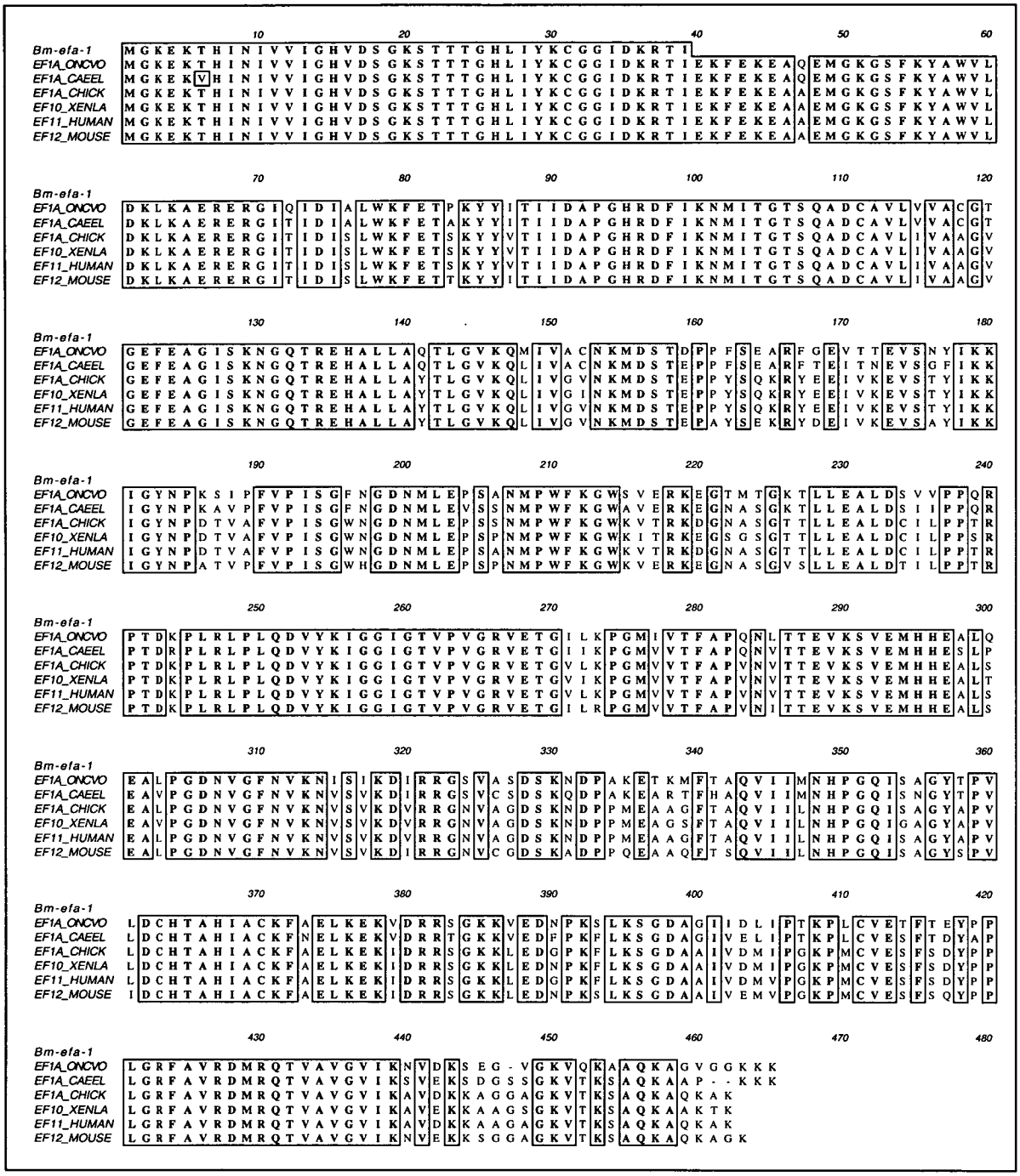


Figure 3.8: Alignment of Bm-EFA-1 with members of the elongation factor-alpha family. *Accession numbers are as follows: Bm-EFA-1; U84736, EFIA_ONCVO;Swiss-Prot P27592, EFIA_CAEEL;Swiss-Prot P27592, EFIA_CHICK;Swiss-Prot Q90835, EFI0_XENLA;Swiss-Prot P13549, EFIA_HUMAN;Swiss-Prot P28347, EFI2_MOUSE;Swiss-Prot P10126.*

3.2.7 150 bp vL3 band : *Bm-rps*-19

Clones produced from the 150 bp band were homologues of the small ribosomal subunit S19. The predicted protein aligned with the first 22 residues of the *Ascaris suum* S19 protein but was truncated by mis-priming of the oligo dT primer (Fig. 3.9). The corresponding section in *A. suum* gene is particularly A-rich at this point. The *A. suum* gene RS19G is one of the genes eliminated by chromosomal diminution, the discarding of genomic DNA from presomatic cells during embryonic development (Etter et al., 1991). A second slightly diverged gene, AS19S, is retained in somatic cells (Etter et al., 1994). The elimination of DNA encoding genes from presomatic cells is thought to mediate a switch in which ribosomes are initially composed of the RS19G product but contain RS19S in somatic cells representing an alternate method of gene regulation.

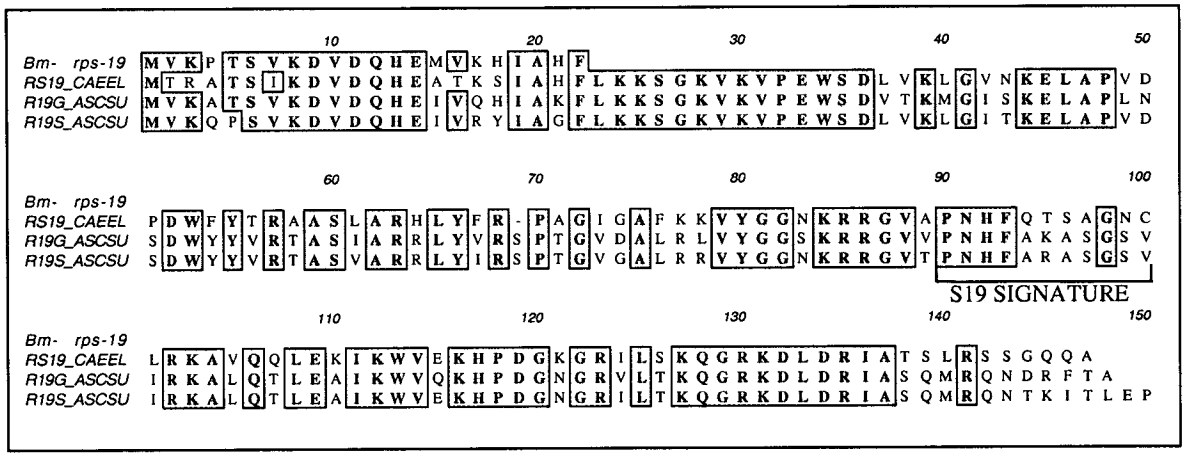


Figure 3.9: Alignment of Bm-RPS-19 with members of the ribosomal protein family S19. *Accession numbers are as follows: Bm-RPS-19, GenBank U80979; RS19_CAEEL, SwissProt O18650; R19G_ASCSU, SwissProt P24494; R19S_ASCSU, SwissProt P39698. The S19 signature is boxed.*

3.2.8 L4 abundant bands : 525 bp is *Bm-cdd*-1

The major component of the abundant L4 525 bp band encodes a homologue of cytidine deaminase (cdd) and has been designated *Bm-cdd*-1. The *B. pahangi* homologue of this gene has previously been cloned (Martin et al., 1996). The predicted protein product of this transcript contains conserved cysteine

residues thought to mediate binding to the catalytic zinc ion (Fig. 3.10). The *B. malayi* and *B. pahangi* sequences differ in only one reside out of 132. Asp-57, a common amino acid at this position in other organisms, is encoded by the *B. malayi* sequence, and serine in *B. pahangi.* To assess whether this is a position of natural variation in the *Brugia* sequences ESTs representing the *B. malayi* sequence were analysed: all ESTs encoded an asparagine at position 57.

Cytosine delaminate was originally identified as part of the pyrimidine salvage pathway via its ability to convert cytidine to uridine. However, recently it has been shown to be the enzyme responsible for producing an alternatively edited form of apolipoprotein B mRNA by converting a cytidine to uridine, thereby forming an in-frame stop codon. Given these functions, the stage-specificity of CDD expression in filariae (our data and (Martin et al., 1996)) is intriguing and may indicate a period of intense gene expression in post-parasitic larvae via its role in a salvage pathway. Recombinant CDD from *B. pahangi* has been shown to bind AU-rich RNA templates (Anant et al., 1997). It was additionally found to bind apolipoprotein B mRNA but failed to edit the transcript. Given the overall high AT content of the filarial genome (Hammond, 1994) this may indicate a general role in regulating the stability of mRNA rather than editing RNA templates.

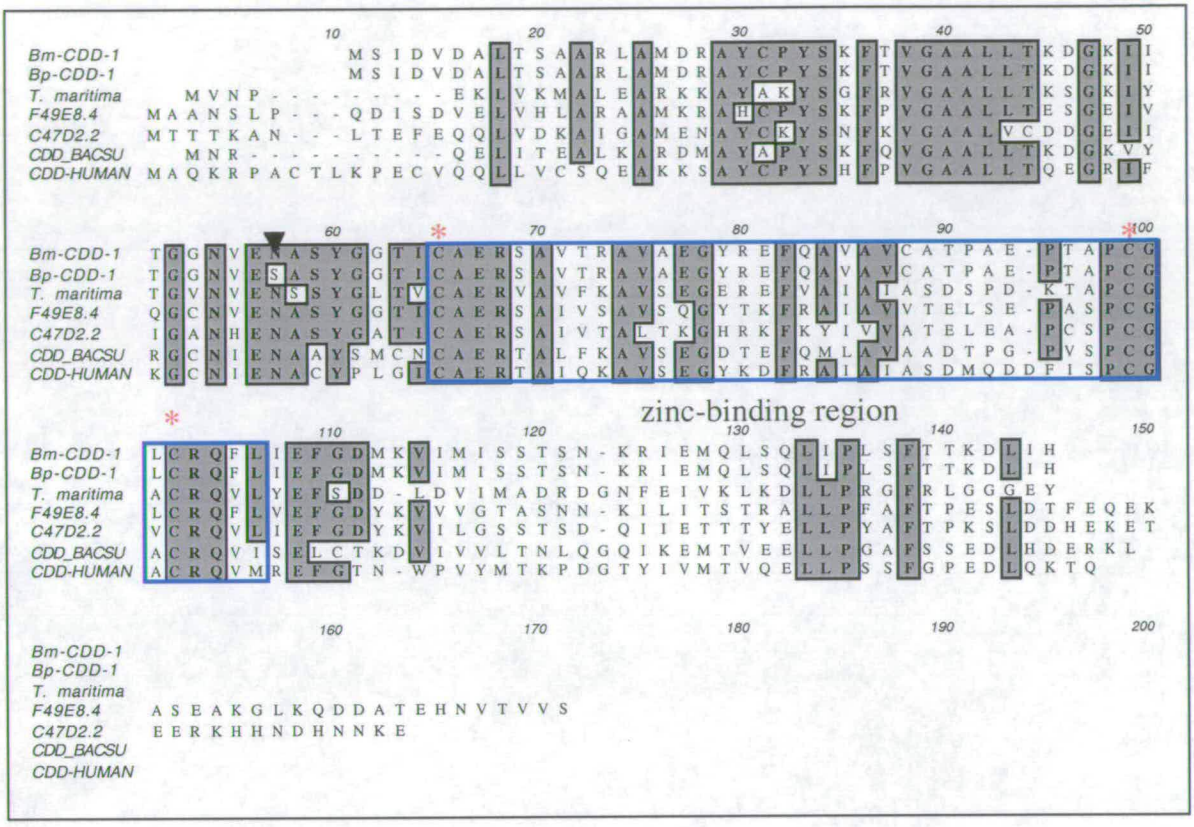


Figure 3.2.10: Alignment of Bm-CDD-1 with cytidine deaminanses. *Accession numbers are as follows: Bm-CDD-1, TrEMBL P90706; Bp-CDD-1, TrEMBL Q93143; Thermotoga maritima, GenBank AE001751; C. elegans cosmid F49E8, U61949; C. elegans cosmid C47D2, U64861; CDD_BACSU, Swiss-Prot P19079; CDD_HUMAN, Swiss-Prot P32320. The zinc-binding region is underlined in blue and the potential coordinating cysteines are marked with an asterix. The potential N-glycosylation site in the B. malayi sequence is arrowed. In the B. pahangi sequence the site is absent as asparagine is replaced with a serine.*

3.2.9 Stage-Specific Expression of Abundant Transcripts

The assignation of *alt-1*, *glt-1*, *cpi-1 and cdd-1* as stage-specific abundant transcripts was confirmed by PCR amplification of these genes from cDNA libraries taken from each stage of the *B. malayi life* cycle. Gene-specific primer pairs were designed to amplify the region of cDNA predicted to encode the mature protein. Each primer pair was used to amplify genomic DNA and cDNA. In each case the genomic product was larger suggesting the presence of introns within each

gene and allowing us to distinguish cDNA products from potential genomic contamination.

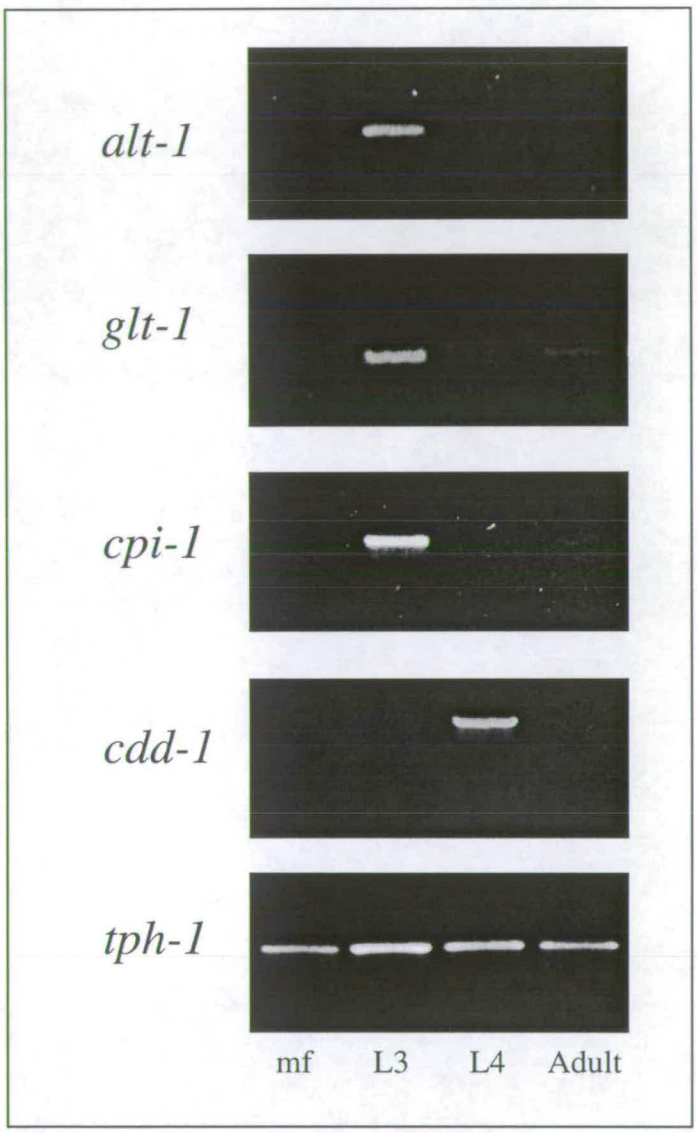


Figure 3.11: *A survey of the stage-specificity of the abundant transcripts using cDNA libraries from microfilariae (mf), infective larvae (L3), day 10 larvae (L4) and adult male (Adult) as templates.*

Using cDNA libraries donated by the Filarial Genome Project as templates, expression of *alt-1*, *glt-1 and cpi-2* is shown to be restricted to the L3 stage (Fig.3.11). In contrast *tph-1* is expressed at comparable levels in all stages, as represented by their cDNA libraries. Identical results were obtained using the same *tph* primers to amplify poly-A primed first strand cDNA from Mf, L3 and mixed-stage adults (Zang et al., 1999). RT-PCR analysis of RNA extracted from L3 harvested from jirds at daily intervals post infection shows that as early as 24 hours after infection RNA levels of *alt-1* and *cpi-2* have fallen below the detection limit of the assay (see Chapters 5 and 4 respectively) indicating tight control of expression, possibly under the influence of mammalian host factors. Although cytidine deaminase appears highly represented in the L4 library, weaker *cdd-1* products are present in all the other libraries.

Some indication of stage-specificity for related transcripts has previously been reported in other filarial species. For example, the *alt-1* homologue in *D. immitis* is expressed only by vector-borne L3 and during the early part of L3 development in mammalian culture conditions (Frank et al., 1996). In *B. pahangi*, *cdd* was identified in a screen designed to select transcripts up regulated in post-parasitic L3s after entry into the mammalian host (Martin et al., 1996). More broadly, the Filarial Genome Project provides data on clone frequency in libraries from each stage of *B. malayi,* which are fully consistent with the results reported here. Thus, *alt-1* has been identified 33 times in randomly picked clones from L3 libraries, but not once from other life cycle stages. Similarly *cpi-1* and *glt-1* have only been identified as ESTs in L3, and *Bm-cdd-1* only in the L4 library, from which it has been sequenced 96 times.

DISCUSSION

Mosquito-stage larvae of lymphatic filariae develop primarily within the fight muscles of their vector species becoming infective to the mammalian host at around 8 - 10 days post infection. At this point larvae (vL3) are large, measuring 1.4 - 1.9 mm in length and must reside within the mosquito until the next blood feed without harming its host. Because transmission to the mammalian host cannot be predicted the larvae must also be ready to respond to a sudden and dramatic change in environment. The thirteen genes identified here may offer clues to this success. Within this small sample of genes are a number of stage-specific transcripts, many of which contain secretory leader peptides and are therefore likely to be localised the cuticle or secreted by the parasite.

The presence of an abundant collagen transcript in vL3s is noteworthy since cuticle synthesis was not thought to occur in the developmentally-arrested L3 stage present in the vector mosquito. Collagen synthesis is also unexpected in day 9 L4s and in day 25 adults, as these parasites have completed the third and fourth moults respectively. Although Howells and Blainey (Howells and Blainey, 1983) have suggested that growth between moults is accomplished largely by stretching a pre-formed, highly annulated cuticle, the abundance of collagen at these time points does suggest a role in growth or thickening of the cuticle between moults. In *C. elegans* it has been demonstrated that collagen genes are transcribed in a carefully controlled temporal cycle, which repeats within each intermoult period. The timing of the *B. malayi* moult is obviously longer (8-9 days between moults rather than 6-10 hours) but expression of specific collagen genes at this early intermoult time may indicate a similar extended patterning in *B. malayi* collagen gene transcription. In addition ESTs representing the *Bm-col-3* gene have been sequenced from day 10 and day 25 post-infection libraries in addition to vector-derived L3's. Larvae at day10 and day 25 post-infection have largely completed the third and fourth moults respectively, further suggesting that significant collagen synthesis occurs after moults, and, that in day 25 parasites this synthesis is not contributing to synthesis of completely new cuticle.

Using small quantities of parasite RNA it has been demonstrated that developmentally regulated larval transcripts can be cloned simply by ligation of gel-purified abundant SL/dT amplified RT-PCR bands from agarose gels. A similar approach has been reported by Yenbutr and Scott (Yenbutr and Scott, 1995), who by isolating a 1.4 kb abundant SL RT/PCR band from vL3 cloned a mixture of transcripts. Two technical points seem worthy of comment. First, in this study less variety was observed within each gel-isolated abundant band when compared with Yenbutr and Scott. This is most likely due to the re-amplification of isolated bands prior to restriction digestion and ligation in our protocol. Secondly, truncated transcripts have been identified due to mis-priming of the oligo-dT primer to A-rich stretches. This is likely to remain a minor limitation to the technique. On the other hand, there appears to be little mis-priming of the SL primer, as the 5' ends of each newly determined sequence showed good alignment with the 5' ends of the homologues deposited in various databases.

It is noteworthy that several transcripts identified repeatedly by EST sequencing of the *B. malayi* L3 library (Williams et al 2000) were not identified in this study. Such transcripts may lack a transpliced leader sequence or be obscured by abundant bands of similar size. Bm-*tpx*-2 is a member of the thioredoxin peroxidase family of antioxidant proteins (Ghosh et al, 1998) and represents 0.85% of ESTs from the L3 library, much lower than those representing *alt*-1 or *alt*-2 (Williams et al, 2000). None of the clones sequenced represented *tpx*-2, which may be due to the *tpx*-2 (714bp) band being obscured by the *tph*-1 (728) band when resolving the SL-dT produced on a gel. Although all clones generated from this band were not sequenced all were positive when screened by PCR using *tph*-specific primers. Resolving the SL-dT products on a larger, higher resolution gel will no doubt allow the identification of more transcripts. Another abundant transcript identified by EST analysis is a member of the ancylostoma secreted protein (*asp*) family (Hawdon et al, 1996). This transcript is not transpliced and would therefore not be identified in SL-dT products.

A highly abundant member of the *alt* family, *alt*-2, was not identified in this study despite being EST sequenced many more time than *alt*-1. The *alt*-1 and *alt*-2 transcripts are similar enough in size that they would not be distinguished by the agarose gel electrophoresis used in this study. It is possible that there were *alt*-2 transcripts in the 550bp band but were not cloned and sequenced because they were hybridisation positive for *alt*-1 due to their high level of homology to each other at the nucleotide level. To avoid this situation more clones generated from each excised band should be sequenced, ensuring variant transcripts of similar size are identified.

The gene-first approach of this study is an important component in identifying proteins of significance in parasitic infections, especially where limited access to parasite material precludes comprehensive developmental and biochemical studies. The data presented here suggest that by using this approach genes of importance in development can be identified, and that there may be an important bias towards these. Of the 7 most abundant transcripts reported here, 4 contain potential hydrophobic signal sequences and 5 are developmentally regulated. As the Filarial Genome Network rapidly expands our inventory of parasite genes, analysis of developmental expression patterns will be an increasingly critical screen for isolating products of most interest in the host-parasite relationship.

CHAPTER 4

B. malayi cysteine protease inhibitors (cystatins), CPI-1 and CPI-2.

INTRODUCTION

The human filarial nematodes *B. malayi* and *W. bancrofti* together account for the majority of human filarial infections (Michael et al., 1996). They are complex multicellular parasites capable of surviving for 10 years or more in their definitive host. During most of this time they reside in the host lymphatic system provoking a vigorous immune response characterised by extraordinarily high levels of IgG4 and IgE antibodies and the presence of specific, cytokine secreting, T cells incapable of proliferation (Maizels et al., 1995). It has been suggested that these parasites must not only protect themselves from attack from the immune system but probably manipulate the immune system in order to survive (Maizels et al., 1993). The cystatin family of cysteine protease inhibitors are candidate molecules with potential for interfering with the immune system. As broad inhibitors of cysteine proteases that have the capacity to bias the antigen processing capacity of antigen presenting cells (APCs). With the identification of a cystatin-type cysteine protease inhibitor among the abundant infective larval cDNAs we sought to characterise its protein product along with a second distinct cystatin identified through EST sequencing.

The cystatin superfamily constitutes a large and diverse group of cysteine protease inhibitors found in the tissues and body fluids of mammals as well as insects, reptiles, plants, fish, and helminths (Barrett, 1987; Barrett et al., 1986; Turk et al., 1997; Turk and Bode, 1991). They can be grouped into three distinct families thought to have evolved from a common ancestor. Stefins or family 1 cystatins are about 100 amino acids in length and lack carbohydrates and disulphide bonds. Cystatins, or family 2 cystatins are slightly larger, disulphide-bonded, secreted molecules. Kininogens are much more complex plasma glycoproteins containing two or three contiguous cystatin domains followed by an additional domain that contains the bradykinin sequence. The precursor molecule is processed releasing the vasoactive kinin peptide.

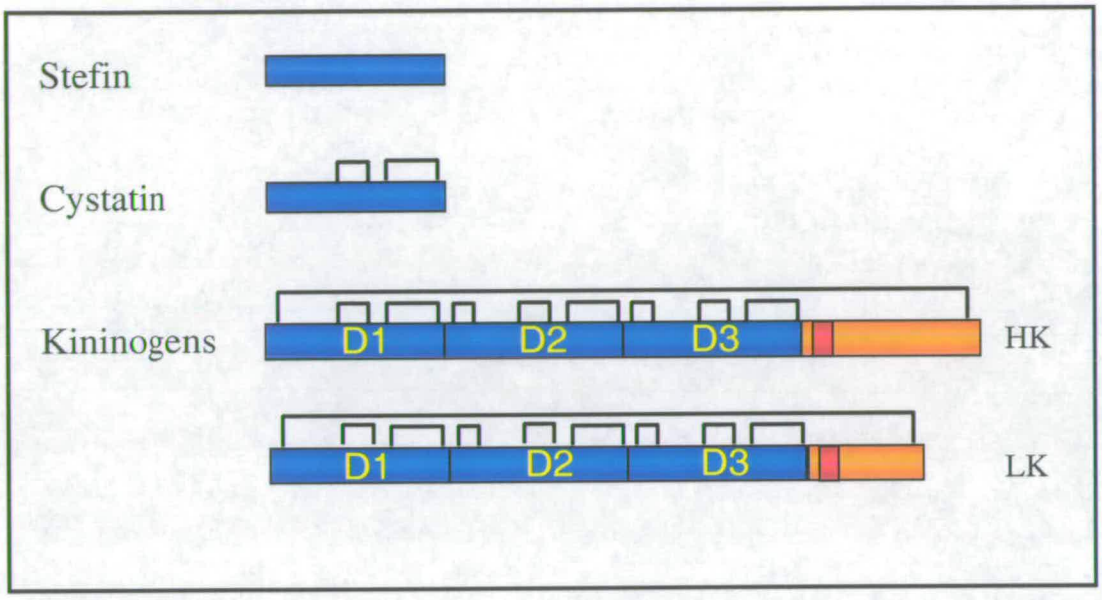


Figure 4.1. Schematic representation of the cystatin superfamily. *Cystatins and the homologous cystatin domains within kininogens are shown in blue. Alternate splicing of the same gene forms the high (HK) and low (LK) molecular weight kininogens. The C-terminal segment (in orange and not shown to scale) contains the kinin segment (in red). Intramolecular disulphide bonds are also shown.*

Possible Roles for Family 1 and 2 Cystatins

The cystatins are well-characterised inhibitors of the papain-like (C1) family of cysteine proteases that includes the lysosomal cathepsins B, C, H, L, S and K that are important mediators of intracellular proteolysis. Recently some cystatins have also been shown to inhibit legumain, a member of the unrelated C13 family of cysteine proteases (Chen et al., 1997). Legumain is an asparaginyl endopeptidase with a strict specificity for the hydrolysis of bonds on the carboxy side of some, but not all, asparagine residues in a protein (Dando et al., 1999), quite different from the bonds hydrolysed by cathepsins. This enzyme has previously been described only in plants and *Schistosoma mansoni*. In *S. mansoni* it is involved in the degradation of host haemoglobin, probably by activating enzymes involved in the proteolysis (Dalton et al., 1995).

A number of physiological functions have been proposed for cystatins. Their role in metastasis has been extensively studied. Some reports have

described increased levels of cystatins in metastatic tumours (Kos et al., 1998; Kuopio et al., 1998). Cystatin genes have also been identified as transcripts down regulated in tumours in comparisons with their normal, matched tissue (Shiraishi et al., 1998; Sotiropoulou et al., 1997). Stable transfection of cystatin C cDNA in melanoma cells led to decreased invasiveness *in vitro* (Coulibaly et al., 1999; Cox et al., 1999). These studies indicate that some tumours use cysteine proteases to invade tissues and that this activity is regulated by cystatins. In perhaps an analogous situation, cystatin C appears to have a critical role in regulation of cathepsins B and L secreted by trophoblast giant cells during invasion of the uterine stroma (Afonso et al., 1997).

Cystatins have been shown to be involved in other protease-dependant processes such as inhibition of viral replication (Aoki et al., 1995; Cimerman et al., 1996; Collins and Grubb, 1998; Garcia et al., 1993; Korant et al., 1985), and bacterial replication (Naito et al., 1995) and resorption of bone by osteoblasts (Lerner et al., 1997). They are also insolubly cross-linked into cornified envelopes, specialised structures cross-linked by transglutaminases that form a rigid barrier over epithelial cells producing a vital barrier for the tissue below (Steinert and Marekov, 1997; Takahashi et al., 1996).

Interestingly, mutations in the two human cystatin genes associated with disease cause specific neuropathologies. One of a number of single point mutations in the stefin B gene is associated with a rare form of inherited epilepsy, known as progressive myoclonic epilepsy of the Unverricht Lundborg type (Pennacchio et al., 1996). Cystatin C deposited in the small blood vessels is found in the brains of Icelanders affected by a genetic disease that leads to fatal cerebral haemorrhages. As well as being truncated at the N terminus by 10 residues the precipitated protein was found to contain a $Gln^{68} \rightarrow$ Leu substitution (Ghiso et al., 1986).

Roles for cystatins in the immune system have also been described. Native, full length, cystatin C has been shown to inhibit phagocytosis by human polymorphonuclear neutrophils (Barna and Kew, 1995). Incubation of mouse

peritoneal macrophage with cystatins, stefins and kininogens has been shown to cause the release of nitric oxide (Verdot et al., 1996).

Functions of Cystatins and Cysteine Proteases in the Immune System

Cysteine proteases participate in the generation of class II-peptide complexes in antigen presenting cells (APCs) at two critical points and it is probable that cystatins contribute to the regulation of this proteolysis. Firstly, the majority of antigens taken up via the endocytic pathway must be degraded in compartments of the endosome/lysosome system before being complexed with MHC class II αβ heterodimers (Watts, 1997). The lysosomal cysteine proteases have been implicated in this degradation on the basis of their localisation in the lysosomes of antigen presenting cells and use of cysteine protease inhibitors (Chapman, 1998; Watts, 1997). Control of this proteolysis may also be important to the outcome of an immune response because cathepsins are also capable of destroying T cell determinants leading to a decreased response to the epitope (Deussing et al., 1998; Vidard et al., 1991). There appears to be considerable redundancy of the proteases generating peptides as APCs from mice lacking either cathepsin B or the aspartyl protease cathepsin D show no impairment in the presentation of several model antigens (Deussing et al., 1998). Legumain/aspariginyl endopeptidase is also capable of digesting antigens for class II presentation (Manoury et al., 1998).

Secondly, loading of class II molecules with foreign peptides cannot take place unless the invariant chain (Ii) is first removed from its peptide binding groove. Ii forms a complex with class II molecules and targets them to the endocytic pathway preventing them from binding to endogenous peptides. However when the complex reaches the endosomes, which are rich in processed exogenous antigens, Ii is sequentially degraded. Several cathepsins have been implicated in this processing with cathepsin S playing a key role, as it is the only protease identified so far capable of fully degrading Ii. LHVS, a specific inhibitor of cathepsin S prevents the complete degradation of Ii resulting in the down regulation of cell-surface class II (Riese et al., 1996). Recently it has been

shown that the developmentally regulated ability of dendritic cells to present class-II peptide complexes on their surface is related to their inability to process the invariant chain. Alterations in expression and localisation of cystatin C probably regulate cathepsin S activity and thus ultimately control the fate of class II molecules (Pierre and Mellman, 1998).

The outcome of an immune response is also affected when cathepsin S is inhibited. Mice treated with LHVS were immunised with either ovalbumin or the T cell independent antigen TNP-ficoll. Antibody levels to ovalbumin, but not TNP-ficol, were decreased in LHVS-treated animals. In a mouse model of pulmonary hypersensitivity, treatment ablated a rise in IgE titres and profoundly blocked eosinophil recruitment to the lung (Riese et al., 1998) and mice genetically deficient in cathepsin S are protected against collagen-induced arthritis (Nakagawa et al., 1999).

The blocking of cathepsin B activity in mice infected with *Leishmania major* has profound effects on the immune response. Susceptible BALB/c mice completely controlled the infection when treated with the inhibitor. The immune response towards the parasite paralleled resistant strains of mice by mounting a Th-1 type response although untreated, susceptible BALB/c mice mount a Th-2 type response to the infection (Maekawa et al., 1998).

Structural Studies

Structurally cystatins and stefins are well-characterised proteins consisting of five anti-parallel β-pleated sheets, which are wrapped around a central α-helix. Three regions of high amino acid conservation are brought together creating a three-point interaction between inhibitor and protease active site cleft (Bode et al., 1985; Stubbs et al., 1990). The contact points correspond to the extremity of the N-terminal region, a central hairpin loop and a further C-terminal loop. The N-terminal section adopts no regular conformation in solution but is thought to interact with residues just outside the active site cleft of target protease after binding (Lindahl et al., 1994). Mutagenesis and deletion of these residues have

demonstrated that this region makes a strong positive contribution to protease binding affinity. The N-terminal 10-12 residues are not only necessary for high binding affinity but can also determine the specificity of interaction with a range of cysteine proteases (Estrada et al., 1998; Hall et al., 1993; Hall et al., 1995; Mason et al., 1998; Shibuya et al., 1995). These studies have demonstrated an important role for the wholly conserved glycine positioned 10-12 residues from the N-terminus. This is thought to confer structural flexible to the N-terminal residues due to its lack of any side chain.

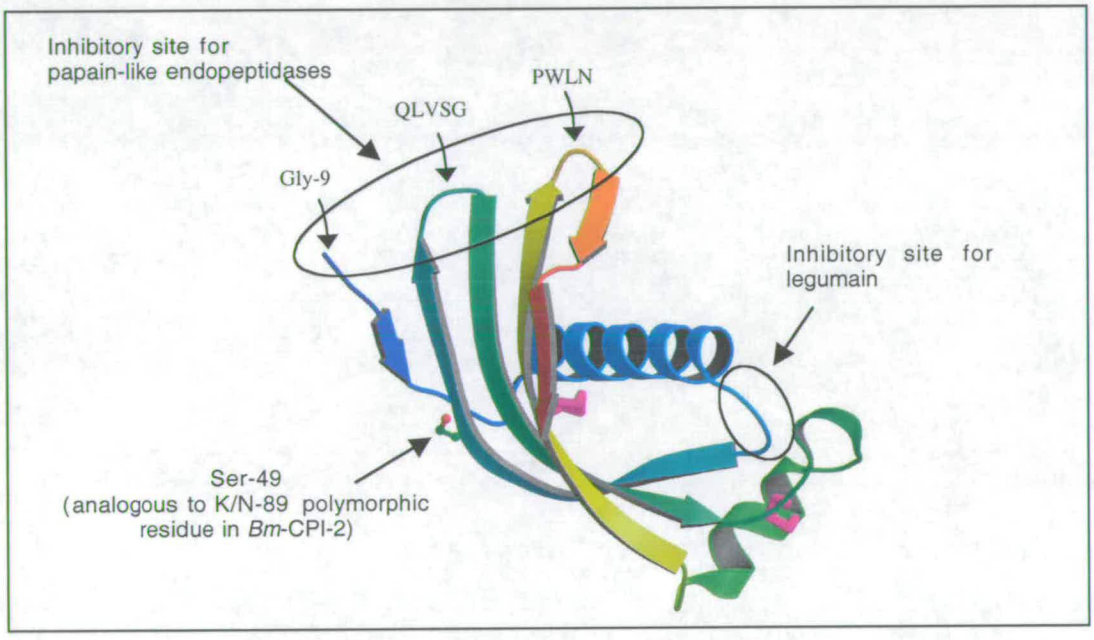


Figure 4.2. Notable features of cyatatin family members:. *Image created in MolScript using co-ordinates for the X-ray crystal structure of chicken cystatin (1CEW). Regions involved in the inhibition of papain-like cysteine endoproteases and legumain are indicated. The proposed legumain-binding motif is only found in a subset of cystatins. The polymorphic residue K/N-89 found in Bm-CPI-2 and onchocystatin sequences aligns with Ser-49 using ClustalW multiple alignment. Polymorphism in the chicken Ser-49 has not been reported. The two chicken disulphide bonds are shown in purple*

The central conserved region which form part of the protease-binding site adopts a hairpin loop structure. The most common sequence at this point is Glu-Val-Val-Ala-Gly although the end residues are the only ones strictly conserved in all members of the superfamily. Mutational studies on this loop have shown that K*i* values can be reduced 10 –1000 fold by changing the conserved glycine residue. The largest decreases in affinity are seen with cathepsin B with inhibition of papain being

similar to the wild type protein and cathepsin L being insensitive to these mutations (Auerswald et al., 1992; Genenger et al., 1991).

Cystatins from Parasites

A wide range of cysteine proteases has been reported from parasitic nematodes (Tort et al., 1999) and it is possible that cystatins serve to control these activities. Specific functions for these proteases are not yet well defined but they can be expected to be involved in host invasion, migration, nutrition, development and perhaps immune evasion.

Cystatins have been reported in other human-infecting filariae, *O. volvulus* and in the rodent filariae *A. viteae*. In *O. volvulus* it has been suggested that onchocystatin controls the activity of cysteine proteases known to be involved in the moulting process (Lustigman, 1993; Richer et al., 1993). The protein has been localised to 'lakes' formed between the old and new cuticles prior to ecdysis (Lustigman et al., 1992). In *A. viteae* the onchocystatin homologue is released by parasites maintained in culture which prompted Hartmann and colleagues to look at the effect of the protein on host cells. They found that the recombinant protein suppressed both polyclonal and antigen-specific proliferation of mouse splenocytes and T cells. The recombinant also induced the release of IL-10, but not IL-4 or IFN-γ from splenocytes (Hartmann et al., 1997b). IL-10 produced by Th-2-like cells in response to *B. malayi* antigen has been shown to have a down-regulatory effect on Th-1-like responses (Mahanty et al., 1997). Sequences of cystatins have also been reported from the sheep parasite, *Haemonchus contortus* (Newlands. and Skuce, unpublished, accession number AF035945) and the free-living nematode *C. elegans* that contains two related cystatin homologues (WormPep R01B10.1 and K08B4.6).

With the identification of one *B. malayi* cystatin gene as a stage-specific gene along with a second distinct cystatin discovered through EST sequencing (Blaxter et al., 1996) we set out to characterise these gene products with the speculation that these proteins are candidate parasite molecules capable of modulating APC and T cell function.

4.2.1 Isolation and Sequence Analysis of *Bm-cpi-1* and *Bm-cpi-2*.

The cystatin gene *Bm-cpi-1* was originally identified by its high level of expression amongst mRNA transcripts from the infective, mosquito-borne *B. malayi* larvae (Chapter 2 and in (Gregory et al., 1997)). This cDNA consists of 526 bp encoding a predicted protein of 127 aa (Fig 4.3). The 5´ end of the gene is *trans*-spliced with the 22-nt nematode spliced leader sequence (SL-1). The SL-1 sequence is followed immediately by a start codon and a region of 19 aa predicted to form a signal peptide (von Heijne, 1986). The calculated molecular weight of the mature protein is 12.1 kDa and it has an estimated pI of 9.34.

A search of the Filarial Genome Project EST dataset in dbEST confirmed the abundance of *cpi-1* (8/3,051 L3 ESTs, 0.3%). This search also identified 13 related cystatin ESTs representing a gene with homology to onchocystatin, which was designated *Bm-cpi-2*. In contrast to the *cpi-1* ESTs which were derived solely from libraries constructed from vector-derived L3 mRNA, the *cpi-2* ESTs were obtained from libraries constructed from post-parasitic (day 6) L3's, adult male and female libraries as well as the L3 library. *Brugia* ESTs representing either *cpi-1* or *cpi-2* were the only sequences with cystatin homology found in the database. One clone, (SW3ICA038, GenBank accession number R47534) containing a 5´spliced leader sequence and a 3´poly A tail and presumed to represent the full-length *cpi-2*, was obtained from the Filarial Genome Consortium as a PCR product generated from amplifying the clone using T7 and T3 vector primers. This was ligated into the T-ended vector pMOS blue (Amersham). Three clones were selected for sequencing and were found to be identical. The insert of 711 bp contained an open reading frame representing 161 aa. The translated sequence is shown in Fig. 4.4.

The *cpi-2* cDNA is similarly *trans*-spliced with SL-1, contains a predicted signal peptide starting 34 bp downstream of the SL-1 sequence, and encodes a mature protein of 136 amino acids (Fig 4.4). The calculated molecular

weight of the mature protein is 15.6 kDa with an estimated pI of 6.51 after removal of the 25 aa signal peptide.

Homology searches using the BLAST algorithm identified both CPI-1 and CPI-2 as most closely related to family 2 cystatins (Fig. 4.5). Mature CPI-1 and -2 are 20% identical at the amino acid level and contain the highly conserved cystatin motif Q-X-V-X-G and the P-W pair (Fig. 4.3 and 4.4). Notably, CPI-2 contains a SND motif, which has recently been implicated in the inhibition of legumains (asparaginyl endopeptidases), by chicken cystatin C (Alvarez-Fernandez et al., 1999). This motif is absent from CPI-1.

Excluding the reported nematode genes, the closest matches to CPI-1 at the protein level are cystatins from the common carp (31.5% identity) and the African puff adder (34% identity). The closest human cystatin homologue is cystatin E/M (31% identity). The closest matches to CPI-2 are the three human salivary cystatins S, SN and SA (28-30% identity). In common with other filarial cystatins (Hartmann et al., 1997b; Lustigman et al., 1991) both *B. malayi* CPI's contain only 2 cysteine residues capable of forming a disulphide bond and consequently lack the second C-terminal cysteine residue pair that spans the P-W residues in all non-filarial family 2 cystatin sequences identified so far (Fig. 4.5)

A noticeable feature of *Bm*-CPI-1 is the absence of an evolutionarily conserved glycine residue near the N-terminus. This was confirmed by sequencing multiple clones and by examining ESTs corresponding to *Bm*-CPI-1. A search of the filarial EST dataset with CPI-1 identified 4 *O. volvulus* ESTs closely related to *Bm*-CPI-1 but distinct from onchocystatin. In common with *Bm-cpi-1* these EST's were derived from a vector-derived L3 stage cDNA library. One clone, SWOvL3CA303 (GenBank accession number R84217), was obtained from the Filarial Genome Project, and the full sequence deposited in the GenBank database (accession number AF177194). The predicted protein, termed *Ov*-CPI-1, is 70% identical to *Bm*-CPI-1 and both proteins have extremely basic PI's (9.3 for *Bm*-CPI-1 and 9.2 for *Ov*-CPI-1). Thus the genes appear to be orthologous. Significantly *Ov*-CPI-1 does encode a conserved N-terminal glycine. Alignment

of the two genes indicates that the divergence of *Onchocerca* and *Brugia* was accompanied by the conservative replacement of a glycine by a serine residue at position 25 in *Bm*-CPI-1 (Figs. 4.3 and 4.6).

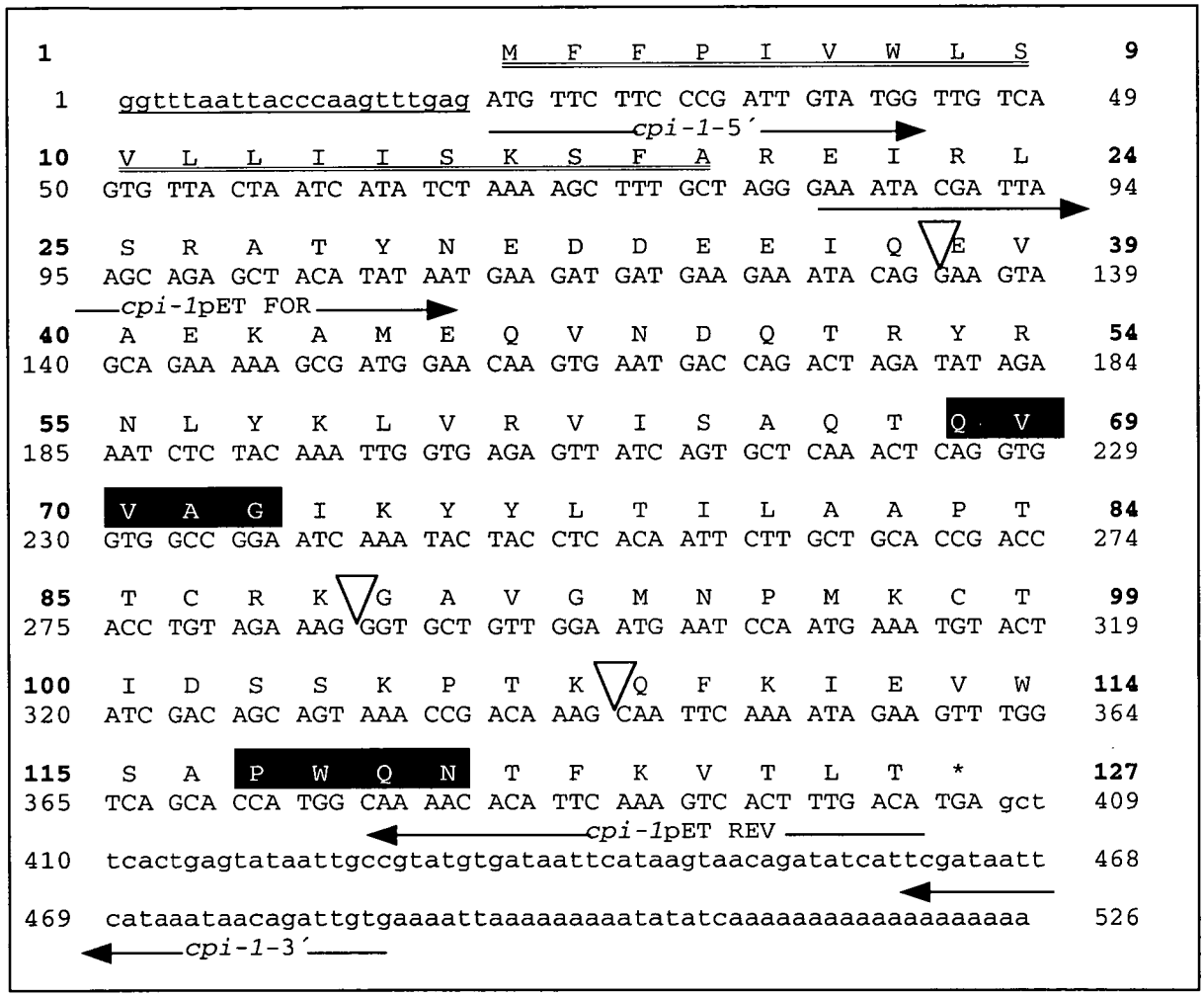


Figure 4.3: Nucleotide and deduced amino acid sequence of Bm-cpi-1. *The 5′ nematode spliced leader sequence SL-1 is underlined. The putative signal sequence is double underlined. Residues thought to be involved in protease binding are solid boxed. Arrowheads indicate the position of introns within the genomic sequence. Primers used for genomic cloning and RT-PCR (5′and 3′) and for expression in pET-29 (pET FOR and REV) are also indicated.*

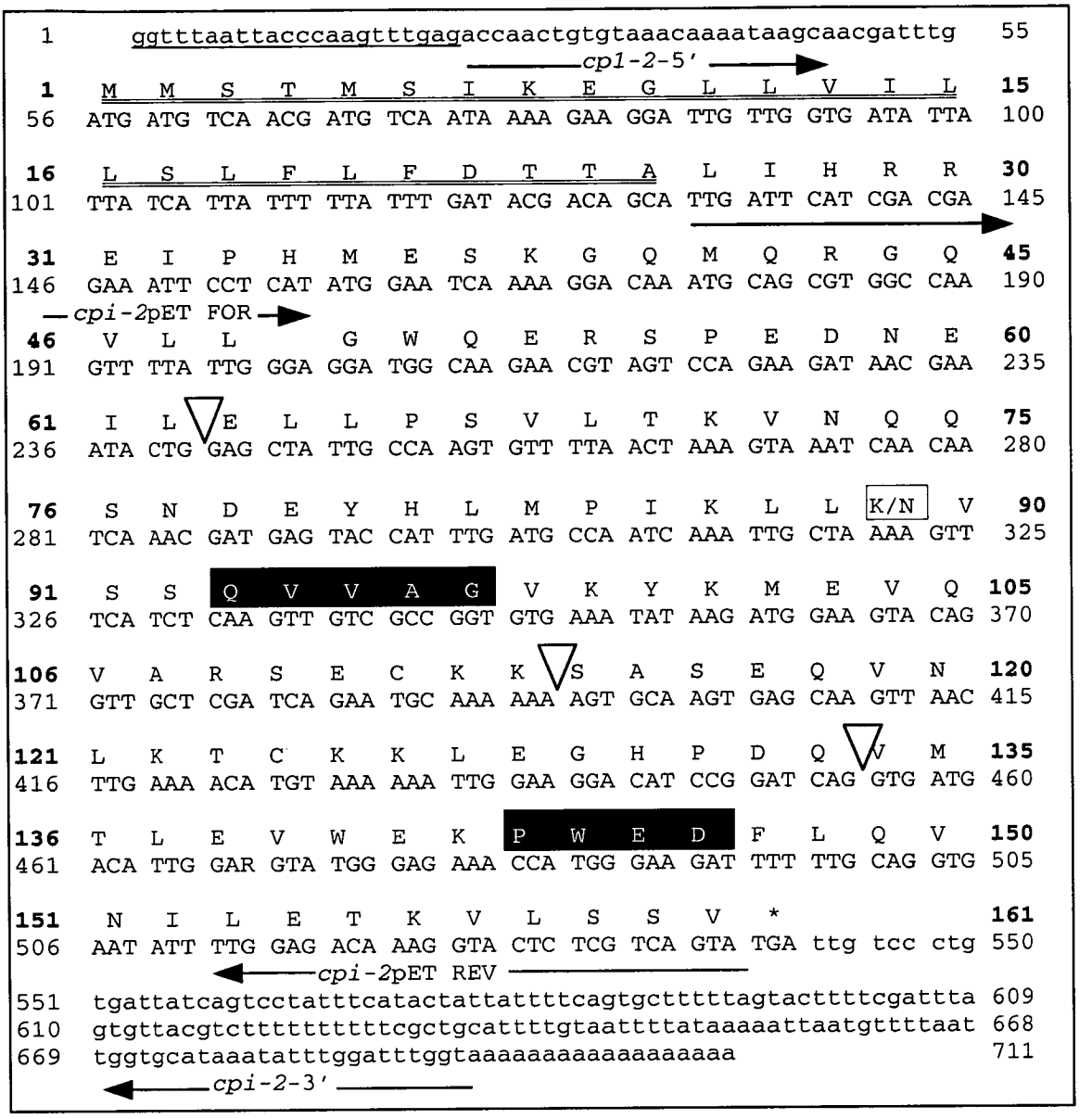


Figure 4.4: Nucleotide and deduced amino acid sequence of Bm-cpi-2. *The 5′ nematode spliced leader sequence SL-1 is underlined. The putative signal sequence is double underlined. Residues thought to be involved in protease binding are solid boxed. Arrowheads indicate the position of introns within the genomic sequence. Primers used for genomic cloning and RT-PCR (5′and 3′) and for expression in pET-29 (pET FOR and REV) are also indicated. The variant residue (K/N-89) is boxed in red.*

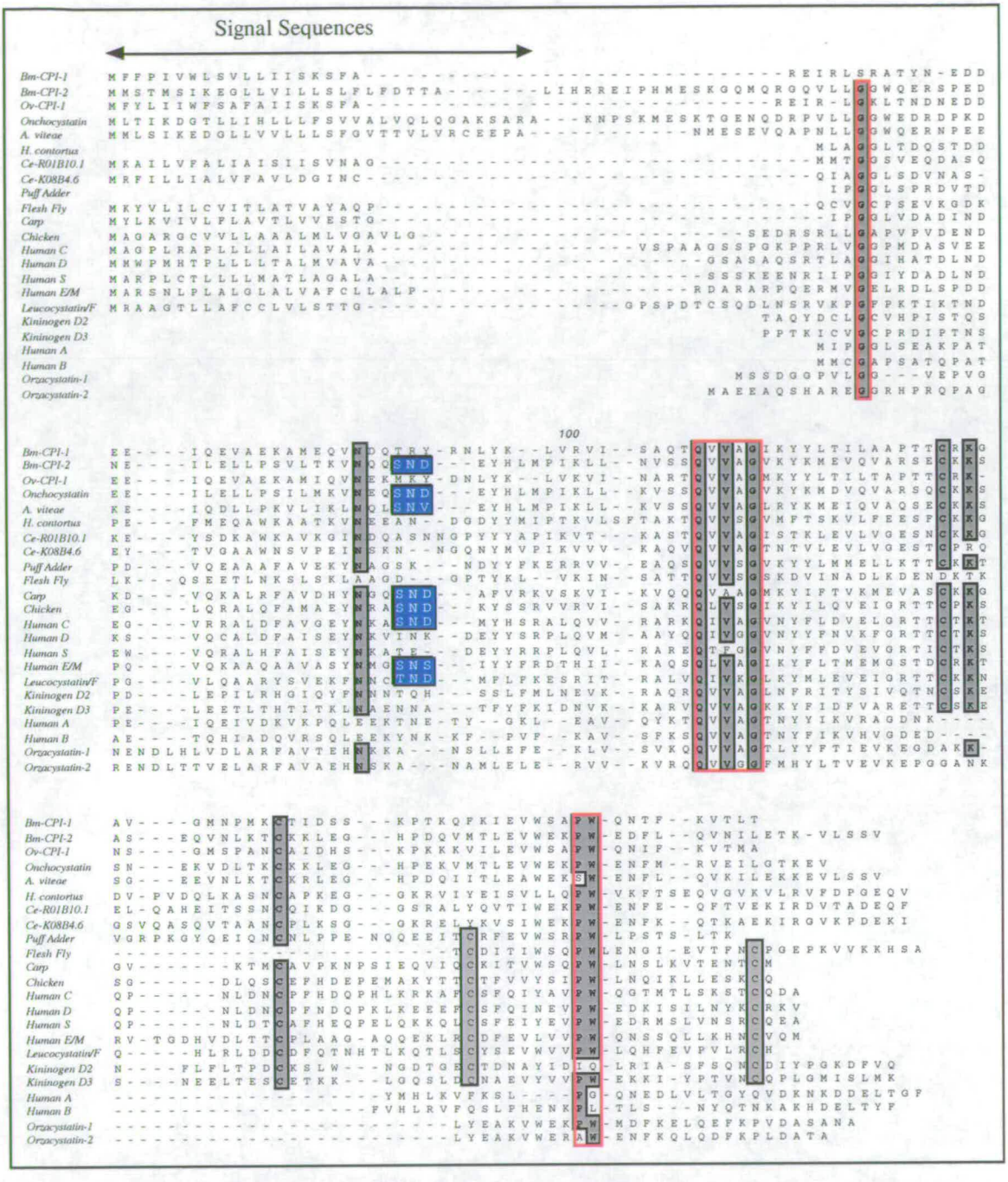


Figure 4.5: Comparison of filarial cystatins with other family 1 and 2 cystatins and kininogen segments. *The alignment was performed using the ClustalW function of MacVector with further manual adjustment to optimise alignment of the highly conserved residues (boxed in red). Residues boxed in blue are those thought to form a consensus sequence for legumain binding.*

Examination of EST sequences in dbEST corresponding to *Bm*-CPI-2 revealed a significant dimorphism: 5/12 sequences code for a lysine at residue 89, with the remaining 7 coding for asparagine. This asparagine is a potential N-glycosylation site as it forms part the consensus Asn-X-Ser/Thr sequence. Deglycosylation of larval extracts with glycopeptidase F did not produce a change in molecular weight of CPI-2 suggesting that the site is not glycosylated by the parasite (Fig, 4.18). An identical dimorphism is observed in ESTs representing onchocystatin from *O. volvulus*. 15/38 sequences code for a lysine residue at the same position, while the remainder code for an asparagine. *Bm*-CPI-1 does not contain potential *N*-glycosylation sites. In *B. malayi* the change in amino acid is achieved by a single nucleotide change (AAA, coding for lysine, and AAC coding for asparagine). Sequencing of CPI-2 inserts in expression vectors have not revealed additional substitutions in the sequence indicating selective pressure on this region in both *B. malayi* and *O. volvulus*.

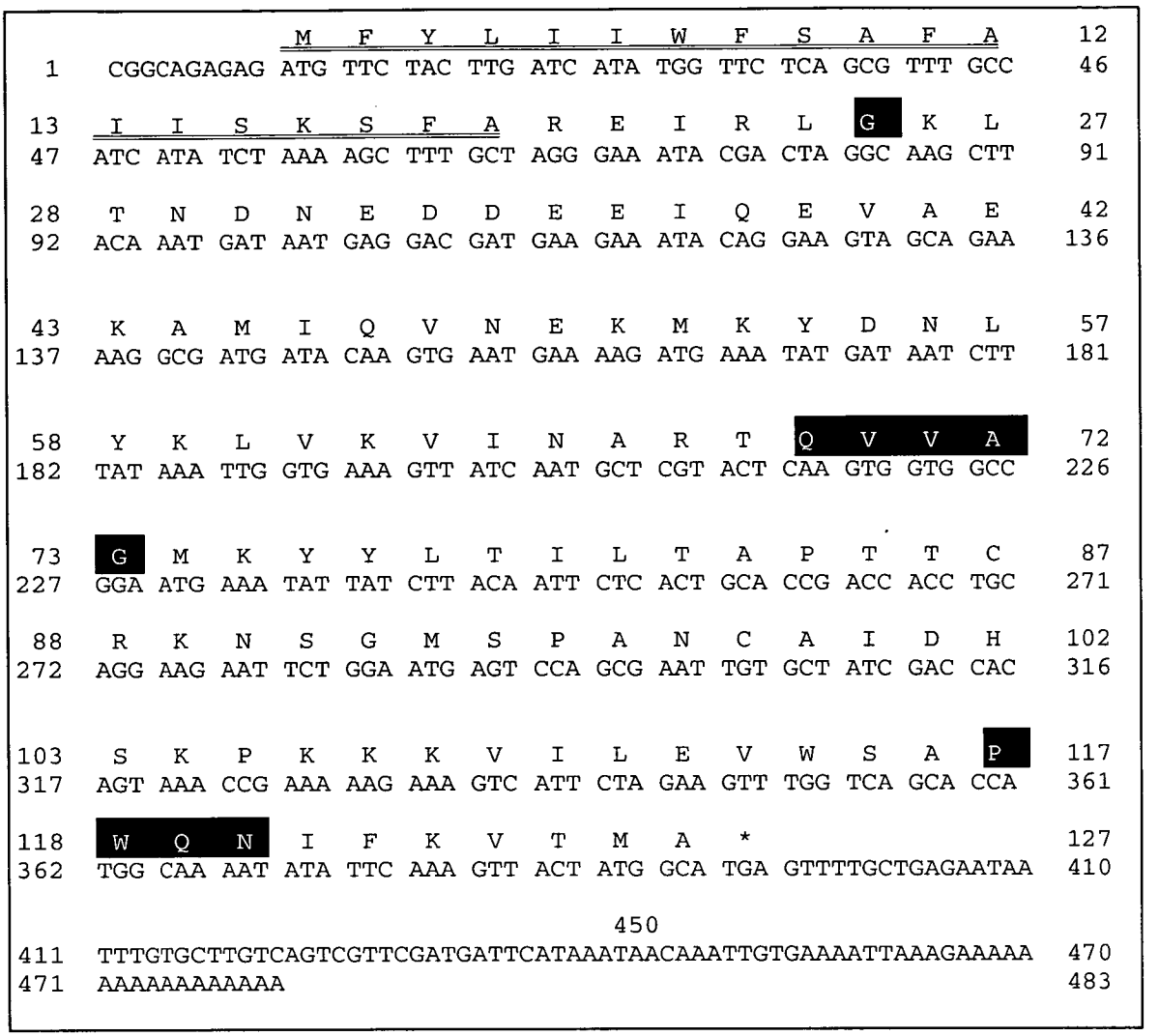


Figure 4.6: Nucleotide and deduced amino acid sequence of Ov-cpi-1. *The 5´ nematode spliced leader sequence SL-1 is underlined. The putative signal sequence is double underlined. Residues thought to be involved in protease binding are solid boxed.*

4.2.2 Analysis of Genomic Structure

A PCR-based strategy was employed to isolate genomic copies of each *cpi* using gene-specific primers situated at the 5´ and 3´ end of each cDNA sequence. PCR carried out using these primers generated products of 1090 bp for *cpi-1* and 1679 bp for *cpi-2* (Figs. 4.7 and 4.8). Both genes contain 3 introns varying in size from 120 bp to 480 bp, which are located at identical positions within the coding sequences (Fig. 4.9). The boundaries of one intron within *cpi-1*

did not match the consensus GT-AG splice donor-acceptor boundaries. The second intron of *cpi-1* uses GC as the 5′-donor site, a sequence known to be very rarely used in *C. elegans* (Blumenthal and Steward, 1997). In a recent survey of *Brugia* genes, a cytosine in the +2 position of the 5′-splice site was only encountered in 2% of *B. malayi* introns (Zang et al., 1999).

4.2.3 Comparison with cystatin genes from *C. elegans*.

Two family 2 cystatin genes have been predicted from the complete genome of the free-living nematode *C. elegans*. Both genes, designated RO1B10.1 and K08B4.6 by the *C. elegans* Genome Project, have been predicted by Genefinder, a program capable of identifying open reading frames and introns from genomic DNA sequence. The program's prediction of introns within the genes is confirmed by comparison with multiple ESTs derived from the genes. These ESTs also demonstrate that both genes are expressed. The two genes are located on different chromosomes: RO1B10.1 is found on chromosome 5 and K08B4.6 is on chromosome 4. The predicted proteins share 48% amino acid identity and have many features in common with cystatins from parasitic nematodes such as the presence of only one pair of cysteines and the exclusive use of QVVAG in the central conserved region. Despite these similarities, the structure of the two genes differs significantly from the *B. malayi* genes (Fig. 4.9). The *C. elegans* genes are interrupted by two introns, only one of which is found in exactly the same position in the *B. malayi* genes.

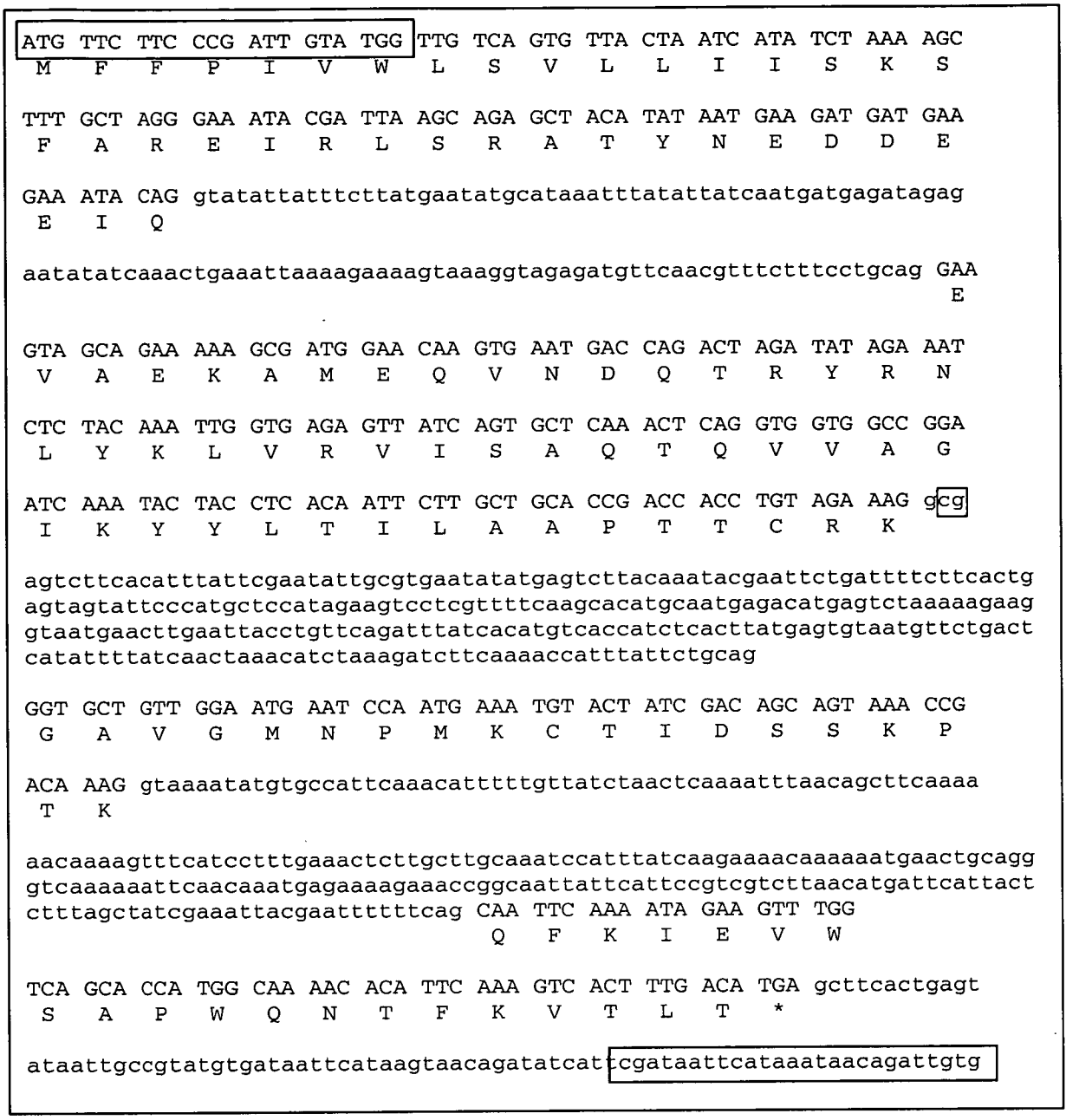


Figure 4.7. ***Partial genomic sequence of the Bm-cpi-1 gene.*** *The nucleotides boxed indicated the gene-specific oligos used to amplify genomic DNA. The Exons are accompanied by the predicted amino acid sequence. The rare GC 5´splice site is boxed in red.*

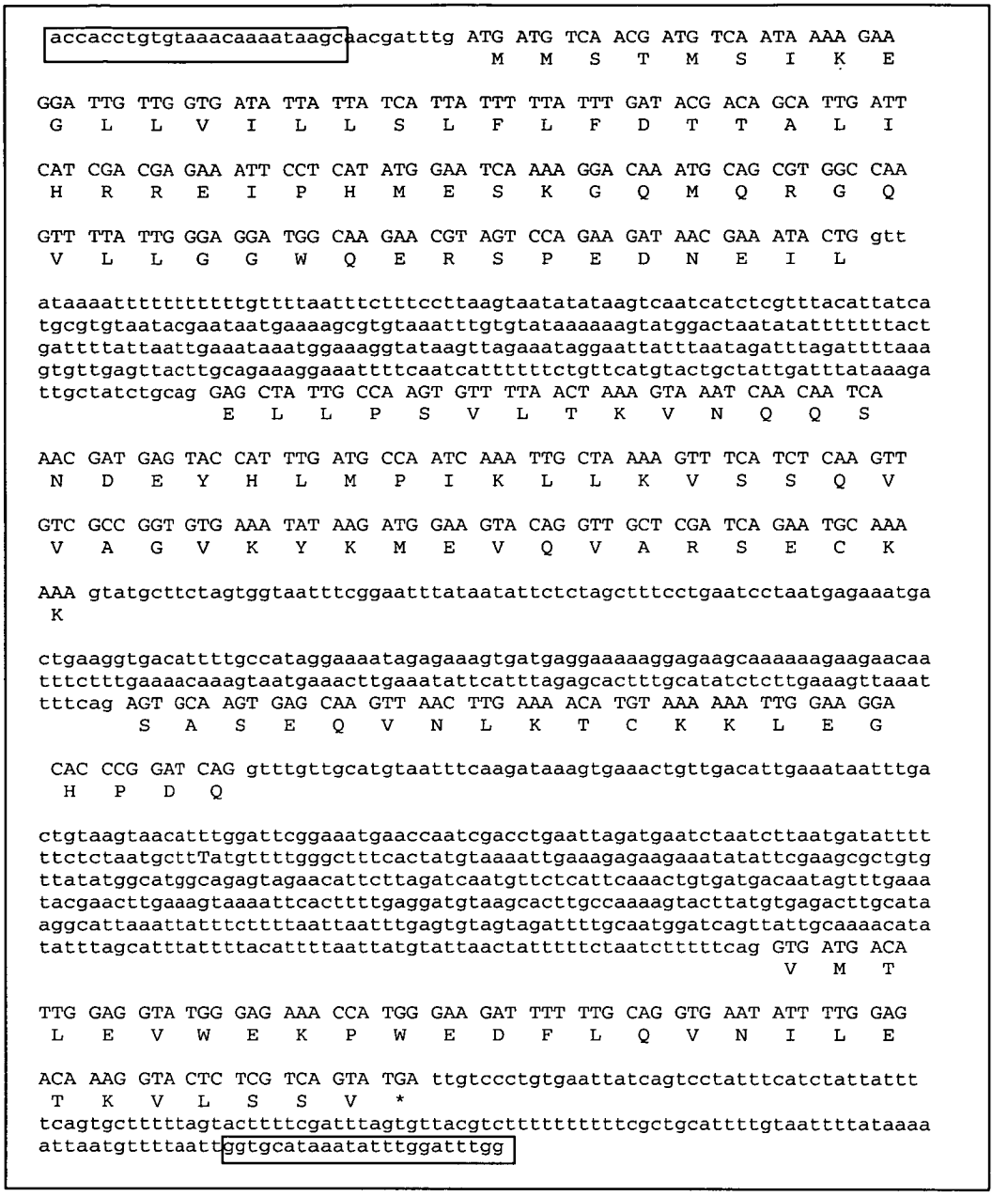
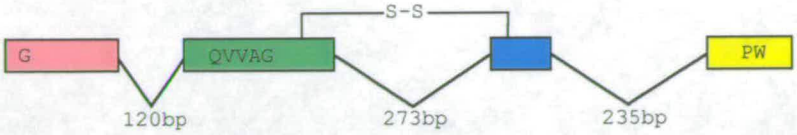


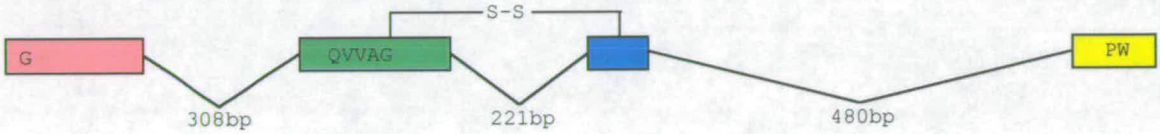
Figure 4.8: Partial genomic sequence of the Bm-cpi-2 gene. *The nucleotides boxed indicate the gene-specific oligos used to amplify genomic DNA. The exons are accompanied by the predicted amino acid translation*

B. malayi

Bm-cpi-*1*



Bm-cpi-*2*



C. elegans

R01B10.1

KO8B4.6

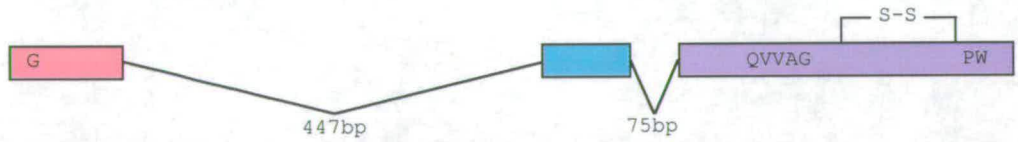


Figure* 4.9:** ***Schematic representation of the gene structures of B. malayi and C. elegans cystatin genes. *Identical box colouring indicates correspondence between exons. Lines represent introns, and their size is indicated below. Conserved residues and proposed disulphide bonds are also shown.*

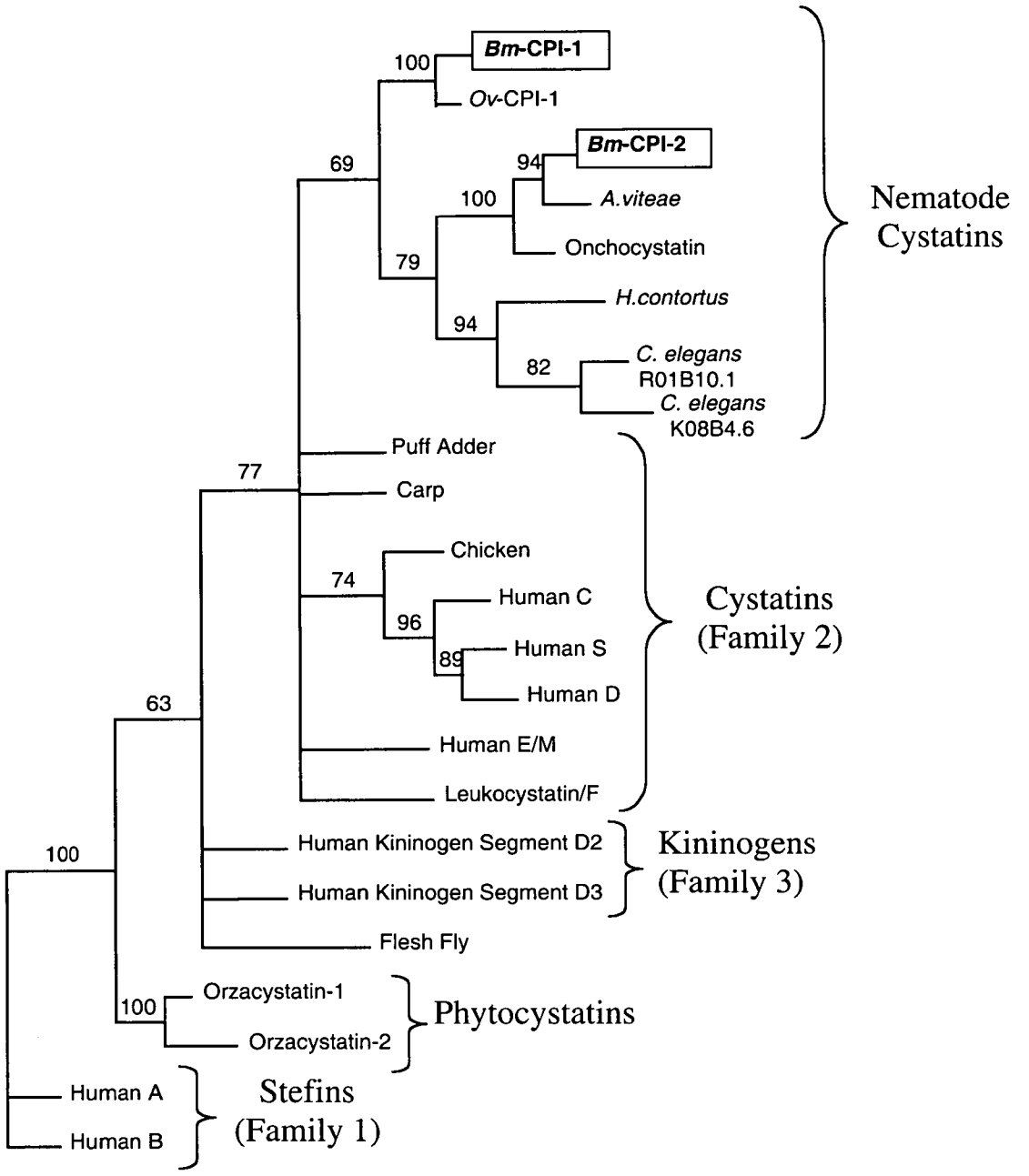


Figure 4.10: Rooted phylogram showing the relationship of nematode cystatins with members of the cystatin, stefin and kininogen families. *Stefin sequences were designated the outgroup for this analysis. Numbers on top of the branch lines show the calculated bootstrap percentage values. GenBank or SwissProt accession numbers are as follows: B. malayi CPI-1; U80972, B. malayi CPI-2, AF015263; A. viteae Av17, L43053; onchocystatin, P22085;Ov-CPI-, AF177194 Haemonchus contortus, AF 035945; C. elegans R01B10.1, AF068718; C. elegans K08B4.6, AF100663; African puff adder, P08935; Carp, P35481; Chicken, P01038; Human C, P01034; Human S, P01036; Human D, P28325; Human M, Q15828; Human Leukocystatin, AF031824; Human kininogen segment 1 (residues 144-207) and Human kininogen segment 2 (residues 266-355), P01042; Flesh fly cystatin, P31727; RICE 1, oryzacystatin-1, P01040; RICE 2, oryzacystatin-2, P20907; Human A, P01040; Human B, P04080*

4.2.4 Phylogenetic Analysis.

Comparison of the available nematode cystatin sequences with representatives from cystatin families 1, 2 and 3 further demonstrated the divergent nature of the nematode cystatins (Fig. 4.6). In addition, *Bm*-CPI-1 and its *O. volvulus* orthologue form a distinct group with no obvious counterparts as yet identified in the complete genome sequence of *C. elegans* (Fig. 4.10).

4.2.5 Analysis of mRNA levels throughout the lifecycle.

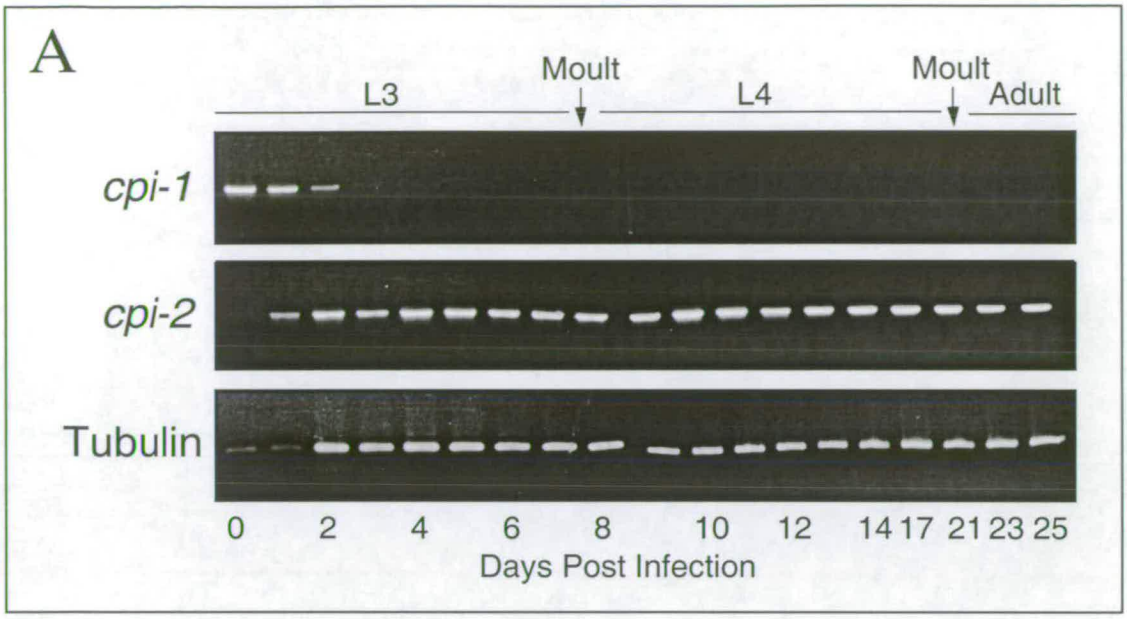
In order to gain information on the transcriptional regulation of the cystatin genes, RT-PCR was employed to detect gene expression throughout the parasite's development in both mosquito and mammalian hosts. Gene-specific primer pairs located within exons 1 and 4 of the genes were chosen to ensure discrimination between cDNA and genomic amplification products. *Bm-cpi*-2 RNA was detected in all life cycle stages tested but *cpi-1* was only detected in the late stages of parasite development in the vector mosquito (Fig 4.11A) and the very early stages of development in the mammal (Fig 4.11B). Twenty-four hours after entry into the mammalian host, *cpi-1* expression fell below the detection limit of the assay, even after 35 rounds of amplification. Very faint bands could be seen at days 5 and 6 post infection and again at days 17, 21, 23 and 25. Control amplifications of the *B. malayi* β-tubulin gene performed in parallel gave a detectable cDNA product at all time points.

Analysis of the distribution of ESTs among the stage-specific libraries used for sequencing also highlights the restricted expression pattern of *cpi-1* and shows that *Ov-cpi-1* is similarly restricted. The distribution of ESTs representing *cpi-2* however indicates an increase in *cpi-2* and onchocystatin expression occurs in the L3 stage (Table 41). This variation around the life cycle is not evident in RT-PCR analysis of *cpi*-2 expression (Fig. 4.11A and B).

	mf	L2	L3	Day 6 L3	Day 9 L3	Adult Female	Adult Male
Bm-cpi-1	0	0	8 (0.3%)	0	0	0	0
Bm-cpi-2	0	0	13 (0.5%)	1 (0.06%)	0	2 (0.05%)	1 (0.02%)

	mf	L2	L3	Moulting L3	Adult Female	ıdult lale
Ov-cpi-1	0	0	4 (1.4%)	0	0	
Onchocystatin	0	0	42 (1.4%)	5 (0.2%)	2 (0.1%)	

Table 4.1. The 'In silico' expression pattern of B. malayi and O. volvulus cystatins as judged by EST abundance. *Numbers of ESTs representing the cystatins are grouped according to their library of origin. The percentage of the total number of ESTs sequenced from each library is also shown. The O. volvulus moulting library was constructed from RNA pooled from L3s after 1, 2 and 3 days of culture. ESTs from libraries generated by subtraction procedures are not included in the analysis.*



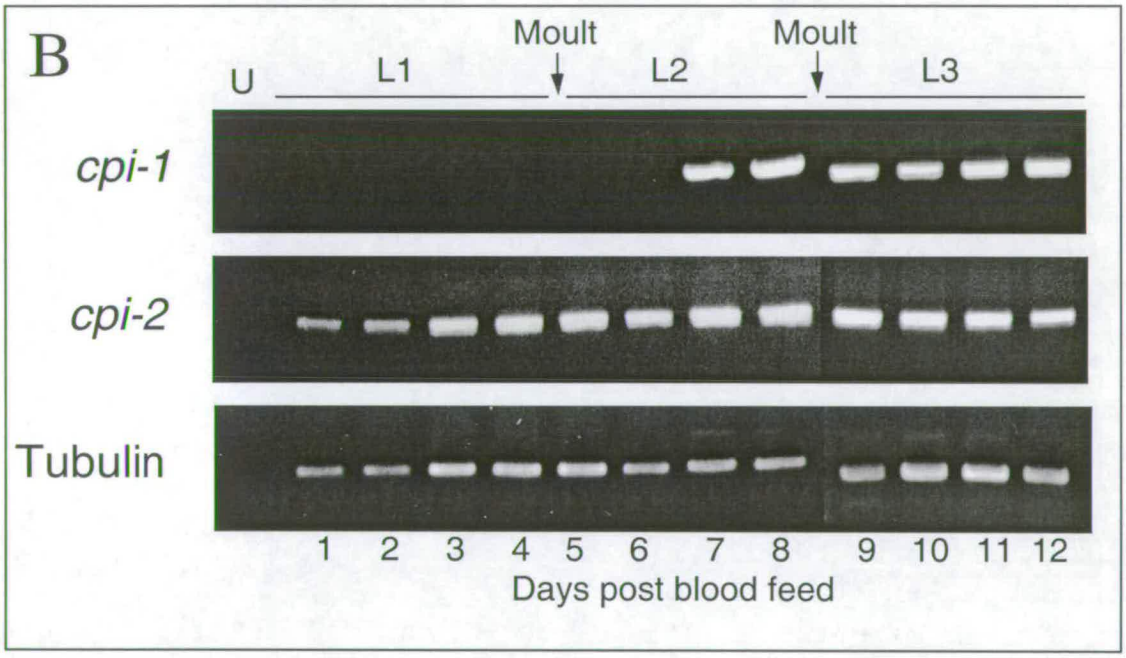


Figure 4.11: Expression of cpi-1 and cpi-2 throughout the life cycle. *The expression of cpi-1 and cpi-2 during development in the mammalian host (A) and the first 25 days of infection of the mosquito (B). U represents RNA extracted from bloodfed mosquitoes fed on uninfected blood. Day 0 represents RNA extracted from vector-derived infective larvae.*

4.2.6. Expression of recombinant CPI-1 and CPI-2.

CPI-1 and CPI-2 have been expressed in both pET-29 and pET-22 expression plasmids. pET-29 (Stratagene) is a T-ended expression plasmid producing fusion proteins with an N-terminal S-Tag and a C-terminal Histidine tag (Fig. 4.12). Inserts are ligated directly into the *EcoR V* restriction site that has been T-tailed by the manufacturer.

Recombinant CPI-2 produced in pET-29 was >80% insoluble without urea present in the sonication buffer. Two prominent bands were present in the urea soluble fraction. CPI-1 expressed in pET-29 was also largely insoluble producing a single band on SDS-PAGE. Both constructs produced sufficient soluble protein to purify recombinants from the Tris-soluble fraction. This allowed antisera to be raised against the recombinant and for enzyme inhibition studies to be carried out.

In an attempt to produce larger quantities of both proteins and to reduce the length of fusion partners both genes were re-expressed in pET-22. This plasmid contains an N-terminal bacterial leader sequence (*pel*B) for potential periplasmic localisation and a C-terminal hexahistidine tag for rapid purification. This construct increases expression of soluble CPI-2 but little improvement was seen with CPI-1. Optimum cleavage of the leader sequence was assumed to occur if the cells were induced at room temperature overnight (Fig. 4.14B). This probably due to the slow rate of production of the fusion protein which allows the *pel*B signal peptidase to cleave more of its target. We have not, however, ruled out the possibility that the lower band is a partially degraded form of CPI-2 although the molecular weight shift is consistent with the removal of the signal sequence.

As CPI-1 lacks the wholly conserved N-terminal glycine residue shown to be essential to the activity of many cystatins a second recombinant CPI-1 was produced in pET-22 lacking 7 amino acids at the N-terminal. The first amino acid of this construct aligns with the first amino acid after the double glycine residues in CPI-2. Fig. 4.13A shows that this construct produced levels of protein similar to the full length CPI-1. Studies using N-terminally truncated cystatins have demonstrated

that these constructs are structurally stable and are useful in assessing the contribution of the N-terminal to the overall cysteine protease affinity (Hall et al., 1995).

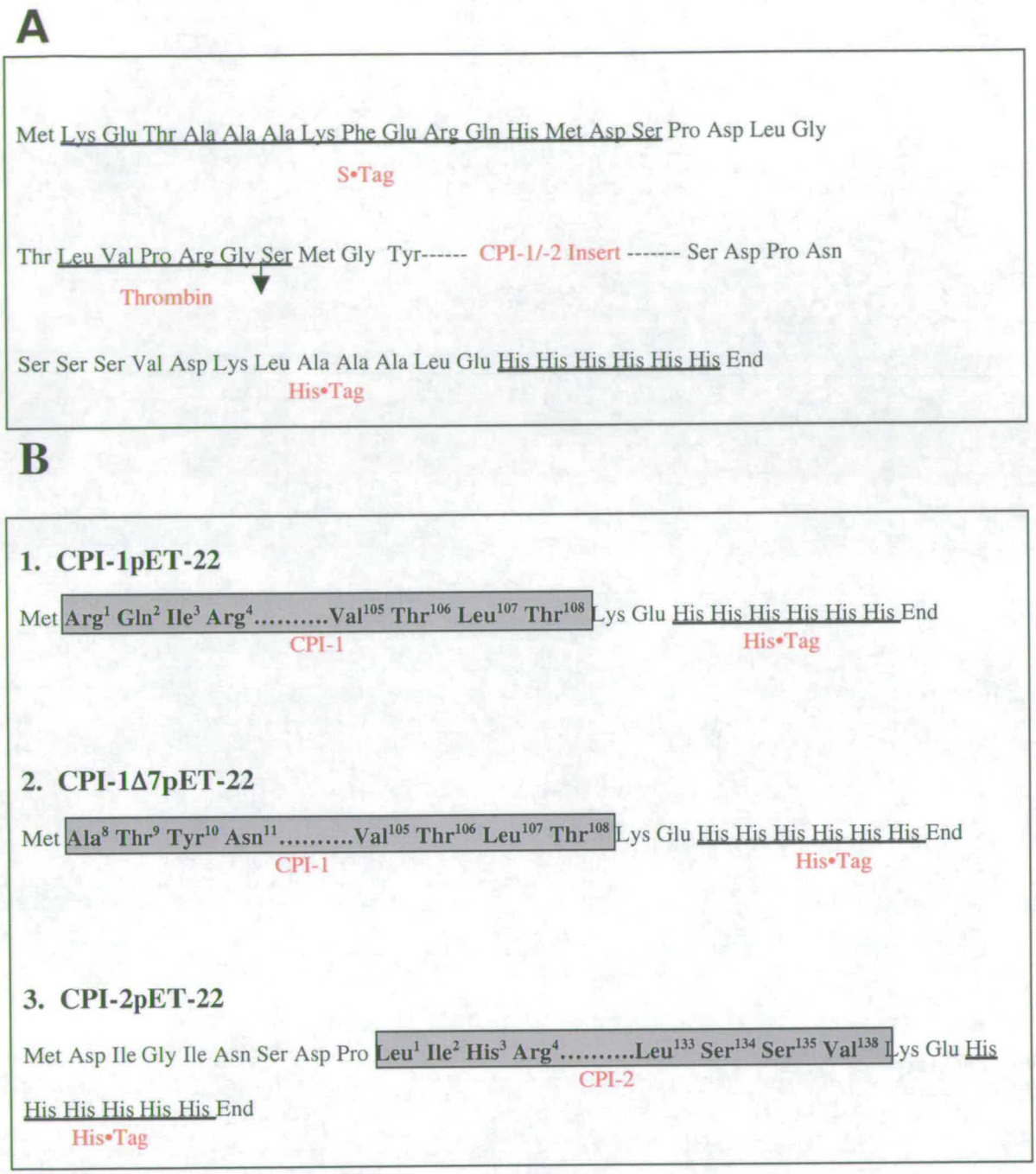


Fig. 4.12: Schematic representations of the recombinant CPI-1 and -2 constructs. Panel A. *Full length, mature CPI-1 and –2 were ligated into the T-ended pET-22 vector.* ***Panel B.*** *CPI-1 and -2 were both expressed in pET-22 as a full length mature proteins In addition a shortened version of CPI-1 lacking 7 N-terminal residues was produced. The first and last residues of the stretch of CPI-1 and –2 expressed are boxed and in bold, other residue are derived from vector sequences.*

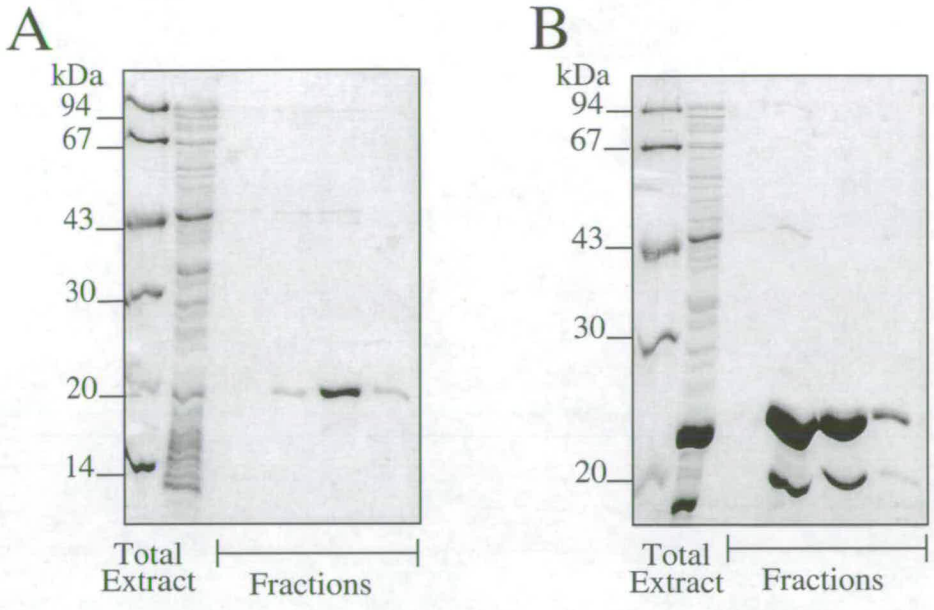


Figure 4.13: Expression and purification of (A) CPI-1 and (B) CPI-2 in pET-29. *Cells were sonicated in His•Bind resin binding buffer and purified over a His•Bind column. Lane 1: molecular weight markers, lane 2: total extracts, lanes 3-6: purified fractions eluted from the column.*

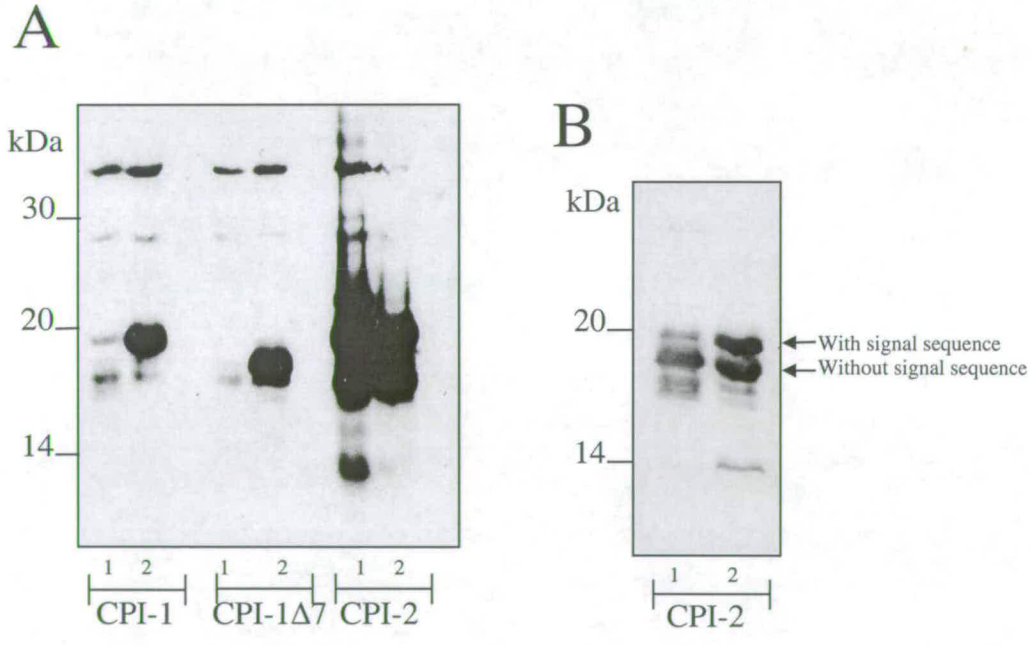


Figure 4.14: Expression of CPI-1, CPI-2 and CPI-1Δ7 in pET 22. Panel A. *Cells were induced for 3 hours at room temperature and extracts were made by sonication in His•Bind resin binding buffer (20mM tris-HCl, 500mM NaCl, 5mM imidazole, pH 7.9). Insoluble material was futher extracted into binding buffer containing 9M urea. Loading for each track represent 200µl of induced culture. A Western blot of the extracts run on a 15% SDS-PAGE gel was probed with a monoclonal anti-histidine tag antibody (Boehringer Mannheim). Lane 1: binding buffer extract, Lane 2: urea extract.* ***Panel B.*** *CPI-2 extracts from cell induced at room temperature to increase the amount of recombinant with cleaved signal sequence. Extracts were prepared and blotted as in panel A.*

4.2.7 Inhibition of cysteine proteases.

The ability of recombinant CPI-1 and CPI-2 to inhibit two distinct cysteine proteases, papain and cathepsin B, was assessed by incubation of the proteases with increasing concentrations of the inhibitors and assessment of the residual activity assayed in the presence of the fluorogenic substrate Z-Phe-Arg-AMC. Both recombinant proteins expressed in the pET-29 system showed significant inhibitory activity towards these proteases, however the inhibitory profiles were distinct for each recombinant protein; CPI-2 showed a preference for papain while CPI-1 was a stronger inhibitor of cathepsin B (Fig. 4.15). As a control, recombinant ALT-1, expressed with identical fusion partner and purified over His•Bind columns, failed to inhibit cleavage of Z-Phe-Arg-AMC by either cysteine protease (Fig. 4.15).

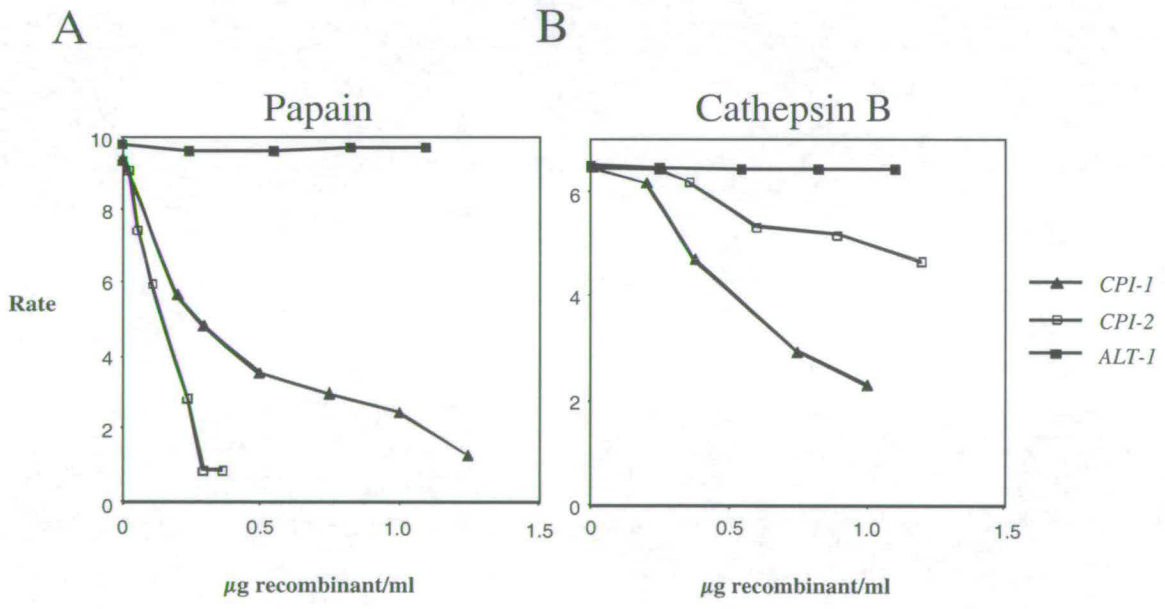


Figure 4.15: Inhibition of (A) papain and (B) cathepsin B by recombinant CPI-1 and CPI-2. *Initial rate of reactions were calculated by monitoring the increase in fluorescence emitted by the cleaved product of Z-Phe-Arg-AMC in the presence of increasing quantities of inhibitor or the control recombinant ALT-1.*

4.2.8 Surface localisation of the CPIs.

Antibodies raised against recombinant CPIs were used to detect the native proteins in extracts of L3 and adults labelled with Iodogen, a surface-specific iodination reagent. Antibodies raised against CPI-1 precipitated a protein of 14 kDa from infective larvae only (Fig. 4.16A), while antisera against CPI-2 precipitated a doublet of 16 and 18 kDa from all life cycle stages (Fig. 4.16B). Results corresponded well with the RT-PCR expression data: while *cpi-1* mRNA and protein could only be detected in larval preparations, *cpi-2* mRNA and protein could be detected in both larvae and adults.

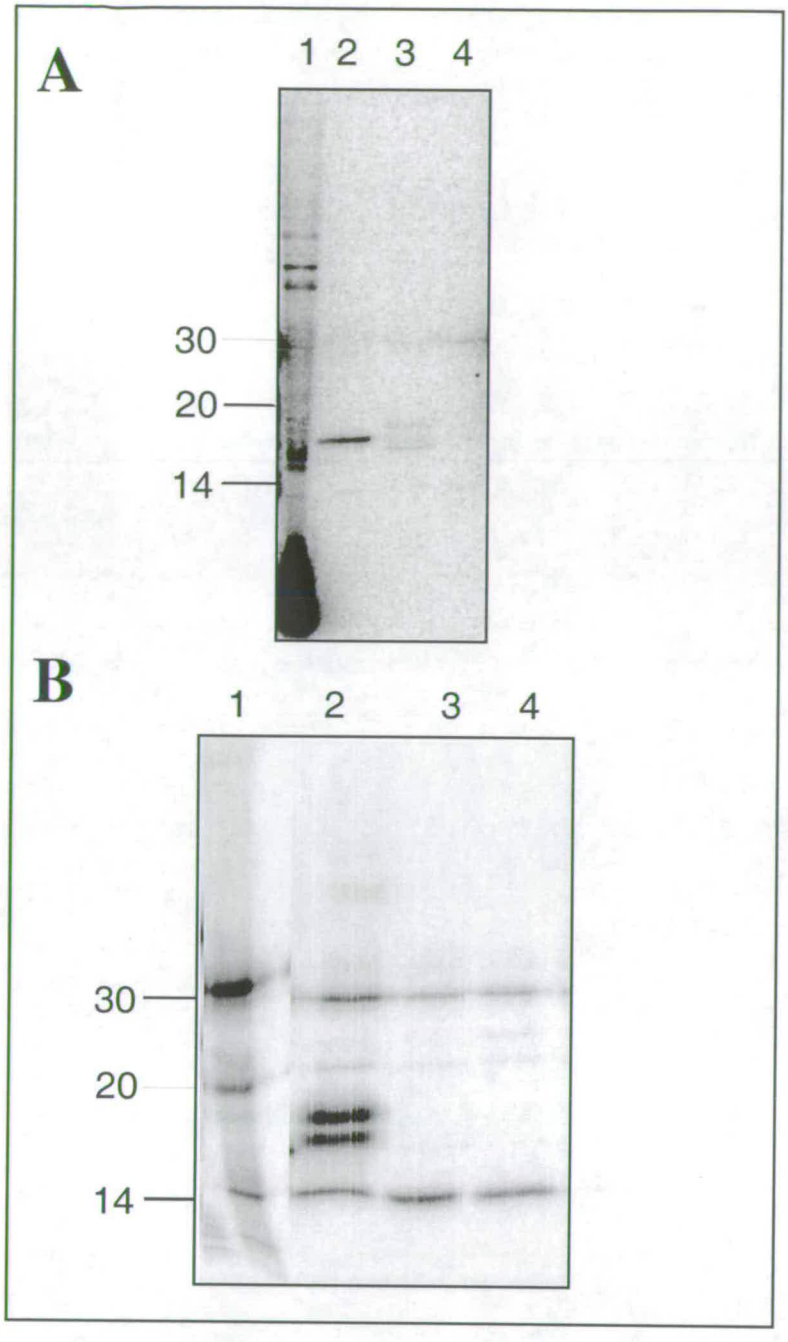


Figure 4.16. Surface localisation of CPI-1 and CPI-2. *Surface componants if infective larvae (A) and adult (B) were iodinated with sodium iodide using IODOGEN as a catalyst. Proteins were solubilised in PBS/1.5% n-octyl glucoside and immunoprecipitated with anti-recombinant antisera.* ***Panel A:*** *Surface labelled L3 extracts (track 1) were immunoprecipitated with antisera raised against CPI-1 (track 2), CPI-2 (track 3) and normal mouse sera (track 4).* ***Panel B:*** *Surface labelled adult extracts (track 1) were immunoprecipitated with antisera raised against CPI-2 (track 2) and CPI-1 (track 3).normal mouse sera (track 4),*

4.2.9 Secretion of CPIs by parasites in culture.

Polyclonal antibodies raised against recombinant CPI-2 were used to immunoprecipitate metabolically labelled proteins secreted by infective larvae and adult worms in culture. Anti-CPI-2 antiserum recognised two bands of 16 and 18 kDa released from infective larvae cultured in RPMI 1640 medium at 37°C (Fig. 4.17). Antibodies raised against recombinant CPI-1 also precipitated a 15 kDa protein (Fig. 4.17).

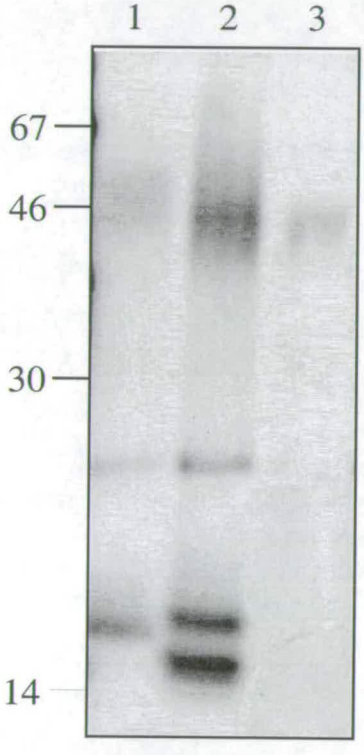


Figure 4.17: Secretion of CPI-1 and CPI-2 by infective larvae maintained in culture.
^{35}S-labelled ES was collected over 3 days and immunoprecipitaed with antisera against CPI-1 (track 1), CPI-2 (track 2) and ALT-1 (track 3).

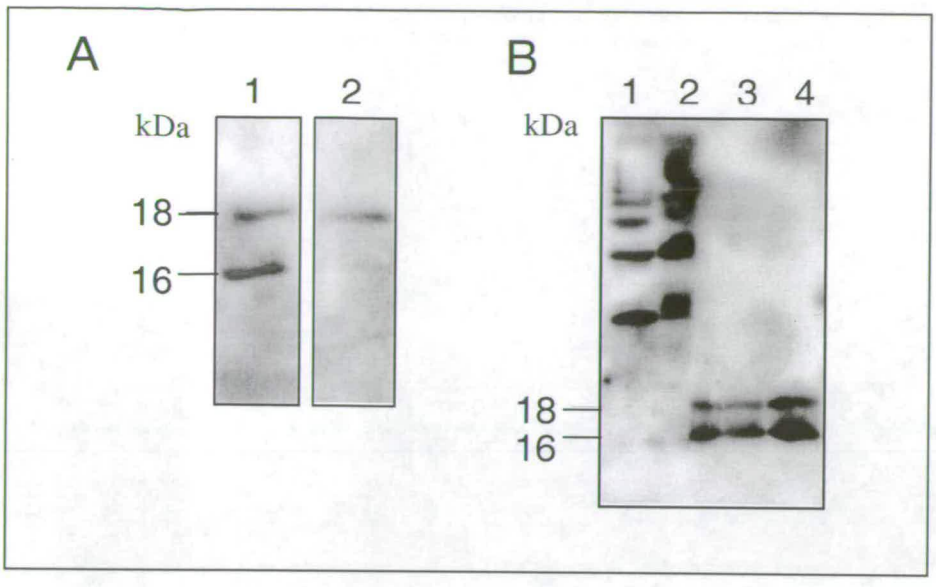


Figure 4.18. CPI-2 resolves as a doublet under reducing conditions and does not contain N-linked glycans. ***Panel A****: Extracts of adult worms were resolved on a 15% polyacrylamide gel in the presence, or absence, of 5% 2-mercaptoethanol in the sample buffer. Both samples were boiled before loading. The gel was transferred to nitrocellulose and probed with antisera to CPI-2.* ***Panel B****: Extracts of adult worms were deglcosylated for 2 hrs with PNGase F (tracks 1 and 3). After separation on a 15% polyacrylamide gel and transfer to nitrocellulose the blot was probed with antisera to either the nematode polyprotein antigen (a gift from Dr. Judi Allen) tracks 1 and 2 or CPI-2 (tracks 3 and 4). Tracks 2 and 4 are negative controls incubated for 2 hrs in PNGase F buffer without enzyme.*

4.2.10 Characterisation of the CPI-2 doublet.

Antisera against CPI-2 recognised a 16/18 kDa doublet in Western blots of adult and larval extracts produced in the presence of a cocktail of protease inhibitors. When these extracts were run in the absence of 2-mercaptoethanol in the sample buffer, only one band of 18kDa was visible (Fig. 14.18A). This result suggests that the cysteine residues in CPI-2 form a disulphide bond and that full denaturation of the protein does not occur under the standard reducing conditions employed. Altered migration under reducing and non-reducing conditions has been shown previously for onchocystatin (Lustigman et al., 1992). Deglycosylation of extracts prior to western blotting failed to alter the migration of the doublet demonstrating that the doublet is not caused by partial glycosylation of CPI-2 and thus the one potential N-glycosylation site is not used (Fig. 14.18B).

DISCUSSION

The cystatins are an ancient and conserved family of cysteine protease inhibitors, and their expression among diverse nematode species is not unexpected. It is curious, however, that cystatins are such prominent gene products among parasitic nematodes, and that *Bm*-CPI-1 has no apparent orthologue in *C. elegans,* suggesting that cystatins play an essential role in transmission, invasion, and/or immune evasion (Hartmann et al., 1997b; Kläger et al., 1999; Lustigman et al., 1992). In *B. malayi*, for example, the cystatins *cpi-1* and *cpi-2* appear to be among the most abundant transcripts from the infective larva, representing 0.5 and 0.3 % respectively of all ESTs sequenced to date. Filarial cystatins are therefore prime candidates for chemotherapeutic or immunological targeting.

Developmental Expression of Cystatins

The distinct developmental expression patterns of the two *cpi* genes may give some insight into their function. The restriction of *cpi*-1 to the mosquito stage indicates that its target protease is likely to be of insect origin rather than mammalian. Initiation of *Bm-cpi*-1 expression coincides with L2 moult and the migration of larvae from the muscle cells of the thorax to the mouthparts of the mosquito (Schacher, 1962a). This is also the time at which larvae become infective to the mammalian host. Expression of the gene is high in comparison with β-tubulin and continues until larvae infect the mammalian host. Localisation of the protein at the surface of the larvae suggests that it may be involved in protection of this structure either while the larvae are resident within the head and mouthparts of the vector or during the early stages of mammalian infection. It is also possible that CPI-1 plays a regulatory role in the extensive cuticular changes that are initiated by the infection process. These changes include altered patterns of surface labelled proteins, biophysical properties and ultrastructural changes.

In contrast, the continuing production and secretion of *Bm*-CPI-2 throughout the mammalian part of the life cycle implies an active role in parasite maintenance. In the related parasite *O. volvulus*, the homologous gene product is

highly expressed in the cuticle of moulting larvae, and indeed synthetic cysteine protease inhibitors are able to arrest moulting altogether (Lustigman, 1993; Richer et al., 1993). However, we saw no detectable increase in *cpi-2* transcription around the moulting events, as for example observed with collagen genes involved in cuticle synthesis in *C. elegans* (Johnstone and Barry, 1996). In addition, the homologous protein from the rodent filarial parasite *A. viteae* has been shown to interfere with host T lymphocyte proliferation (Hartmann et al., 1997b), while in preliminary experiments we have shown that *Bm*-CPI-2 can interfere with MHC class II-mediated antigen processing in human B cells, presumably by blocking host cysteine protease function (B. Manoury, C. Watts, W.F. Gregory, R.M. Maizels, manuscript in preparation). The probability remains therefore that CPI-2, at least, continues to be important in the host-parasite relationship after the final moult to the adult stage.

Protein Structure

Amino acid sequences of cystatin family members contain three features of high homology: (1) an invariant glycine residue within the first 10-15 residues of the mature protein, (2) the central Glu-X-Val-X-Gly motif and (3) a C-terminal Pro-Trp pair. Structural analysis has shown that these regions come together to form a protruding wedge which slots into the active site cleft of the protease (Bode et al., 1988; Stubbs et al., 1990). Exchange of the glycine residue with other amino acids has demonstrated that this residue functions as a hinge between the flexible N-terminal segment and the rest of the molecule. Replacement of the glycine with residues having small side chains results in dramatic reductions in affinity for a range of cysteine proteases.

The replacement of the conserved N-terminal glycine with serine is therefore a notable feature of *Bm*-CPI-1. In mammalian cystatins, mutations in residues up to and including the glycine decrease or abolish the ability of the inhibitor to interact with a variety of substrates (Hall et al., 1993; Hall et al., 1995; Mason et al., 1998; Shibuya et al., 1995). Although serine has a small side chain, other studies have shown that replacement of glycine with serine leads to a significant decrease in inhibitory activity. However, the glycine residue of cystatin

C is not thought to be necessary for inhibition of the aminopeptidase cathepsin C (dipeptidyl peptidase I) (Hall et al., 1993). In addition, truncation of the N-terminal residues of cystatin C has no effect on cathepsin C or cathepsin H inhibition (Abrahamson et al., 1991; Abrahamson et al., 1987). These studies suggest that a cathepsin C-like enzyme may be the natural ligand for CPI-1.

Recently, a second inhibitory site in chicken cystatin C has been proposed which may bind to legumain or asparaginyl endopeptidase enzymes (Alvarez-Fernandez et al., 1999). It is interesting to note that there is sequence conservation between cystatin C and CPI-2 in this loop (SNDMYH and SNDEYH respectively). We have preliminary evidence that CPI-2 is able to block the activity of mammalian legumain and the related enzyme, asparaginyl endopeptidase in antigen-presenting cells (Manoury, B., Watts, C., Gregory, W.F. and Maizels, R.M., manuscript in preparation). This enzyme-inhibitor relationship may therefore be of ancient phylogenetic origin.

Evolutionary Analysis

The set of nematode cystatins now form a robust branch of the superfamily. Interestingly, all lack the second pair of cysteine residues found in family 2 cystatins, residues that clamp the carboxy terminus to the β-pleated sheet in all other family 2 cystatins. Members of the family 2 cystatins are considered to have evolved from family 1 stefins, which lack cysteine residues, acquiring four cysteine residues in the process (Brown and Dziegielewska, 1997; Muller-Esterl et al., 1985; Rawlings and Barrett, 1990). With the view that disulphide bonds are gained but seldom lost during evolution, nematode cystatins appear to be a fixed intermediate in this evolution.

The arrangement of introns in the nematode cystatin genes is also worthy of comment. Most cystatin genes have two introns largely conserved between mammalian family-2 cystatins and cystatin-like segments of kininogens. To date only soyacystatin and leukocystatin genes are known to comprise of 4 exons (Halfon et al., 1998; Misaka et al., 1996) and their overall gene structures are highly

dissimilar to the *Bm-cpi* genes. Despite a relatively close evolutionary relationship and a high level of coding sequence conservation, only one intron position is shared between *B. malayi* and *C. elegans*.

Overall, the *Bm-cpi* genes are atypical in containing three introns, and it is especially surprising that their positions are strictly conserved between two genes whose coding regions are only 25% identical. Thus, the genomic organisation implies a divergence between *cpi-1* and *cpi-2* subsequent to the separation of filarial and *C. elegans* lineages, while the sequence phylogeny suggests the reverse. This may indicate that *Bm*-CPI-1 has undergone rapid sequence evolution from a more recent time-point possibly as an adaptation to the parasitic life cycle.

CHAPTER 5

The abundant larval transcript (*alt*) family of genes

5.1 Introduction.

5.2 Results

- 5.2.1 Abundance of *alt* mRNA throughout the life cycle
- 5.2.2 Other *B. malayi alt* sequences identified by EST
- 5.2.3 Analysis of *alt*-1 and -2 genomic structures
- 5.2.4 Polymorphism within the *alt-2* gene
- 5.2.5 An *alt*-like gene in *C. elegans*
- 5.2.6 Stage-specific gene expression
- 5.2.7 Expression of recombinant ALT
- 5.2.8 Human recognition
- 5.2.9 Protective immunisation

5.3 Discussion

INTRODUCTION

To date the only effective anti-nematode vaccines have been based upon the protective effects of immunisation with irradiated infective larvae. With low level irradiation larvae fail to undergo complete development dying within days or weeks of infection depending on the dose of radiation. During this period they continue to be metabolically active. In contrast immunisation with larval extracts fail to protect against a subsequent challenge infection. This implies that developing infective larvae are the source of parasite antigens capable of inducing protective immunity. It has been suggested that irradiated preparations are effective because they inhibit development and allow a degree of tissue penetration thus increasing the length of time the immune system is exposed to early larval stage antigens and perhaps by presenting protective antigens to a different subset of immune effector cells. Immunisation with material released by infective larvae during *in vitro* culture is also protective. Infections terminated with drug treatment are also capable of inducing significant levels of protection comparable with irradiated preparations. Analysis of the antibody responses of dogs protected against *D. immitis* by a drug-attenuated infection identified a 20/22 kDa doublet in parasite secretions recognised by protected dogs but not by a control cohort undergoing a normal infection. Protein sequencing of the doublet showed both bands to be almost identical and allowed the cloning of the gene. Database searches showed the gene to be novel, although a weak similarity with phospholipase A2 enzymes from honeybees was noted. To date homologues of this gene have only been described in filarial parasites.

Products, such as the *Dirofilaria* 20/22 kDa proteins, which are associated with parasite invasion, which are tightly regulated and parasite-specific, are likely to be essential to the success of parasitism (Bianco et al., 1990; Bianco et al., 1995). In this chapter a family of related, stage-specific genes from *B. malayi* that are prime candidates for a new vaccine against filarial infection are described. These genes, identified by cloning of abundant L3 transcripts ((Gregory et al., 1997) and Chapter 3) and through analysis of EST's (Blaxter et al., 1996) have been designated abundant larval transcripts (*alt*) are close relatives of the *D. immitis* 20/22 kDa

5.2.1 Abundance of *alt* mRNA throughout the life cycle

The *alt*-1 gene (Figure 5.1) was originally identified as a prominent *trans*-spliced mRNA from L3 larvae of *B. malayi* (Gregory et al., 1997), and *alt*-2 as a closely related expressed sequence tag (EST) from the same stage. Three representative *alt*-2 clones were obtained from the Filarial Genome Project and their inserts sequenced in full (Figure 5.2). Since then, the Filarial Genome Project has deposited over 16,000 expressed sequence tags from all stages of this parasite (The Filarial Genome Project, 1999). Analysis of this database reveals that *alt-1* is represented by 33/2378 ESTs (1.39%) from the L3, and *alt*-2 by 67/2378 (2.82%). The *alt*-2 and *alt*-1 transcripts are, respectively, first and third in abundance among all L3-expressed cDNAs in the EST dataset (Williams et al., 2000). Remarkably, neither cDNA can be found among the 14,000-plus sequences derived from other points in the life cycle, implying that *alt*-1/-2 are expressed at less than 0.01% of mRNAs in non-L3 stages.

5.2.2 Other *B. malayi alt* sequences identified by EST sequencing

The searching of over 21,000 *B. malayi* ESTs deposited in dbEST has allowed the identification of additional members of the *alt* family characterised by highly variable, acidic N-terminal regions with conserved central and C-terminal regions. Six further *alt* genes were tentatively identified and clones representing two of these genes were obtained from the Filarial Genome Project and sequenced in full. An alignment of the predicted protein sequences of these new ALT genes with ALT sequences from other filariae is shown in Fig. 5.5.

products. In addition members of the *alt* family have been recognised in other filariae.

In *O. volvulus* two highly similar genes have been described. Ov-ALT-1 was immunoscreened from a L3 expression library using a pool of sera from putatively immune individuals (Joeseph et al., 1998). A second gene was identified through EST sequencing: Ov-ALT-2 shares 85% amino acid identity with Ov-ALT-1. Sequence from an 18 kDa protein secreted by developing larvae was used to clone the *A. viteae* homologue. More recently, a gene encoding an *A. viteae alt* has been reported and its protein product shown to be secreted around the time of the L3 – L4 moult (Pogonoka et al., 1999).

This chapter describes the *B. malayi alt*-1 and *alt*-2 genes in detail and report the presence further members of the family by analysis of ESTs. We have characterised the stage-specific transcription of *alt*-1 and -2 and analysed the genomic structure of the genes. In addition we show that vaccination with recombinant ALT-1 can match the levels of protection achieved by irradiated larvae. The *alt* gene products represent attractive vaccine antigens for three reasons: (i) they are larval-specific in immunological terms; (ii) they are highly expressed, offering an abundant target; and (iii) they have no known homologue in the mammalian host.

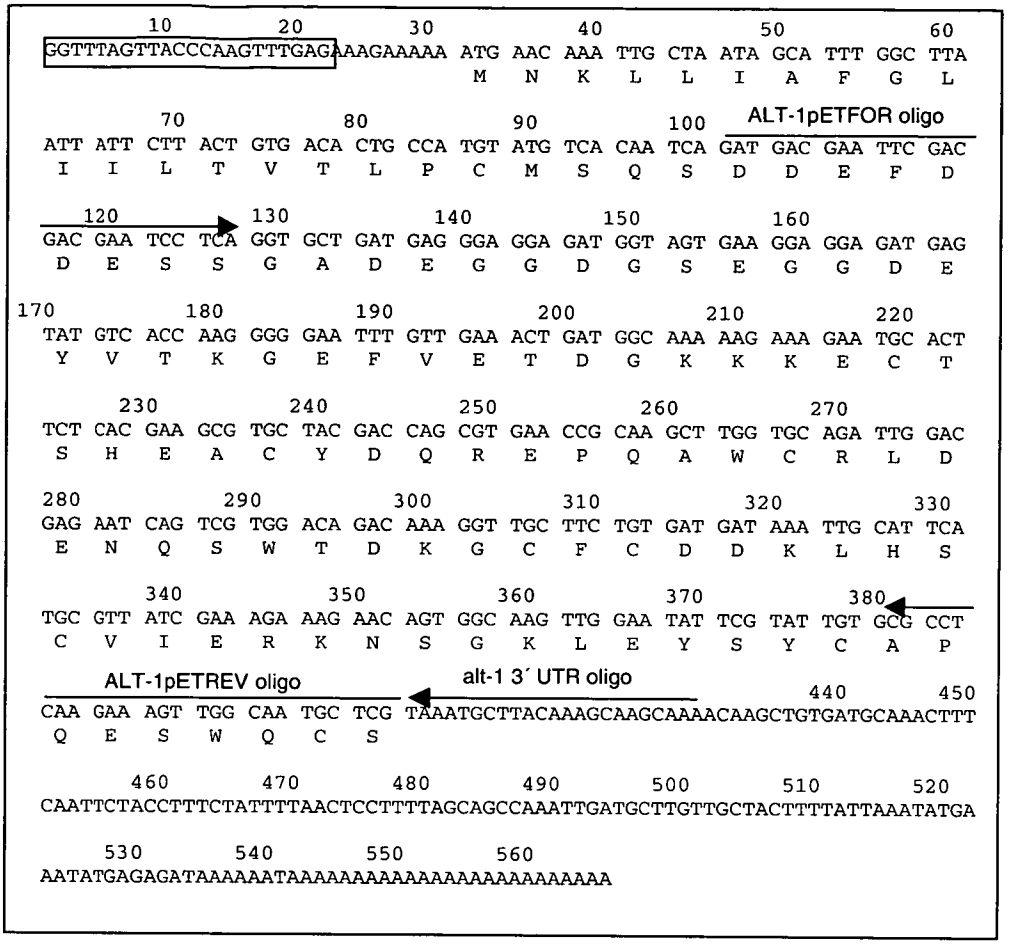


Figure 5.1: Nucleotide and deduced amino acid sequence of Bm-alt-1: *The nematode spliced leader sequence SL1 is boxed and the positions of primers used for expression (ALT-1pETFOR and ALT-1pETREV), PCR of genomic DNA and RT-PCR (ALT-1pETFOR and alt-13´UTR).*

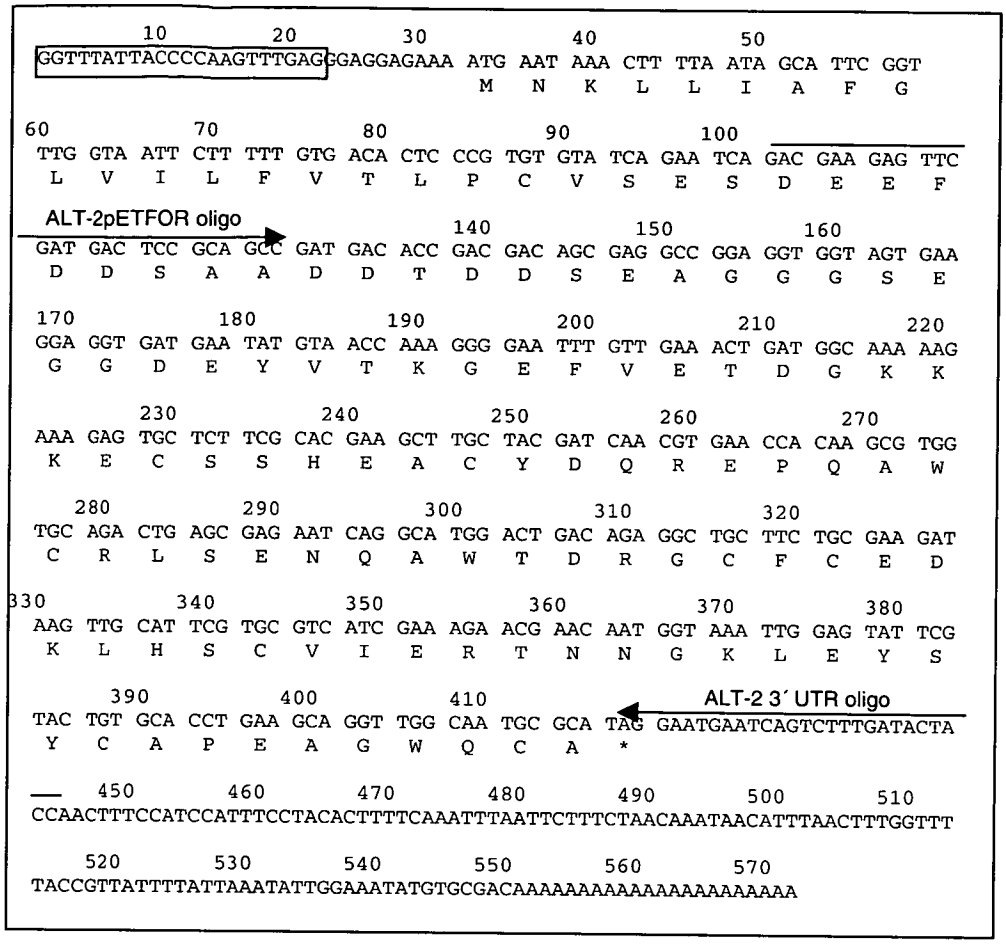


Figure 5.2: Nucleotide and deduced amino acid sequence of Bm-alt-2: *The nematode spliced leader sequence SL1 is boxed and the positions of primers used for expression (ALT-2pETFOR and ALT-2pETREV), PCR of genomic DNA and RT-PCR (ALT-2pETFOR and alt-2 3´UTR).*

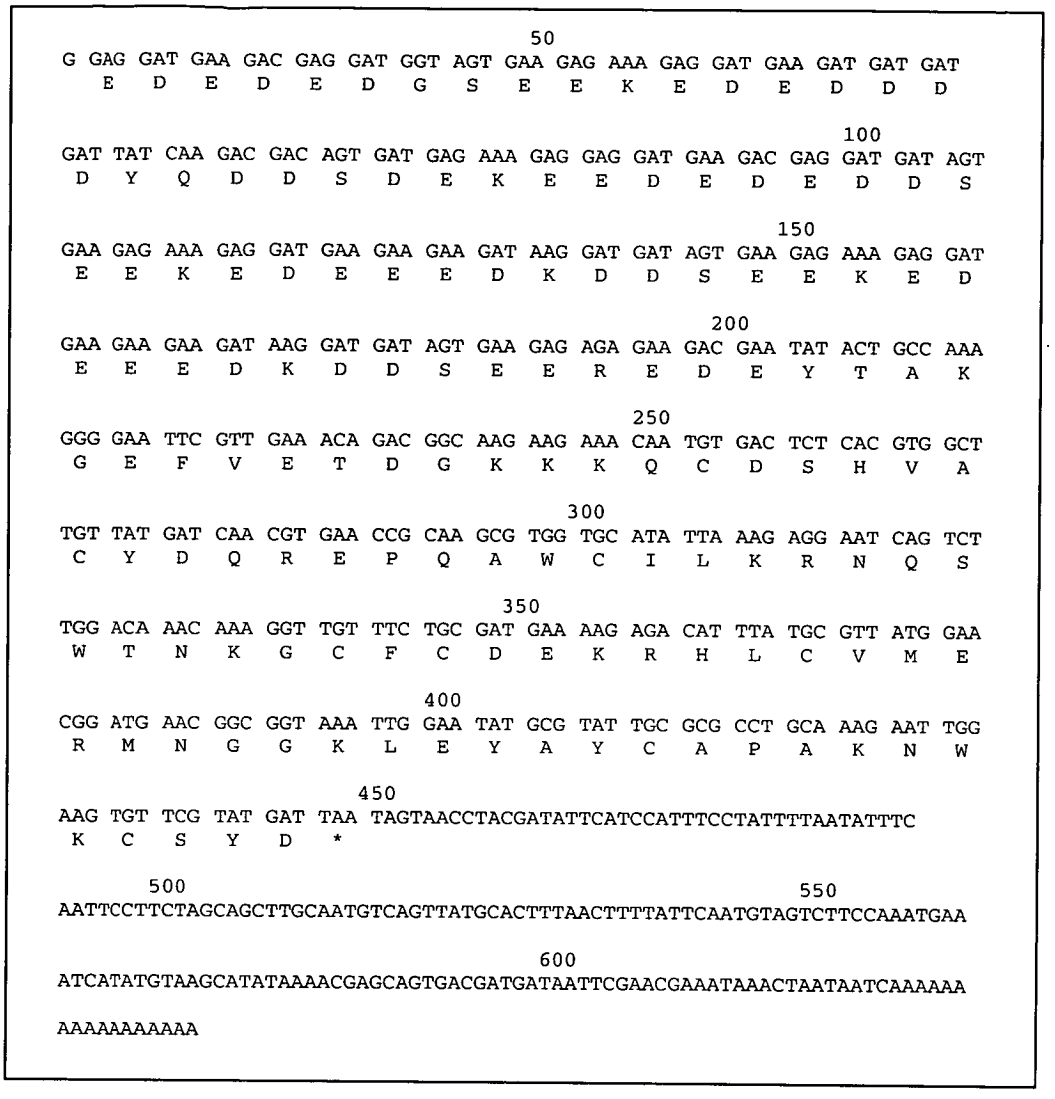


Figure 5.3: Nucleotide sequence and deduced amino acid sequence of Bm-alt-5. *This represents the confirmed sequence of the expressed sequence tag SW3D9CA340SK (GenBank accession number AA585623)*

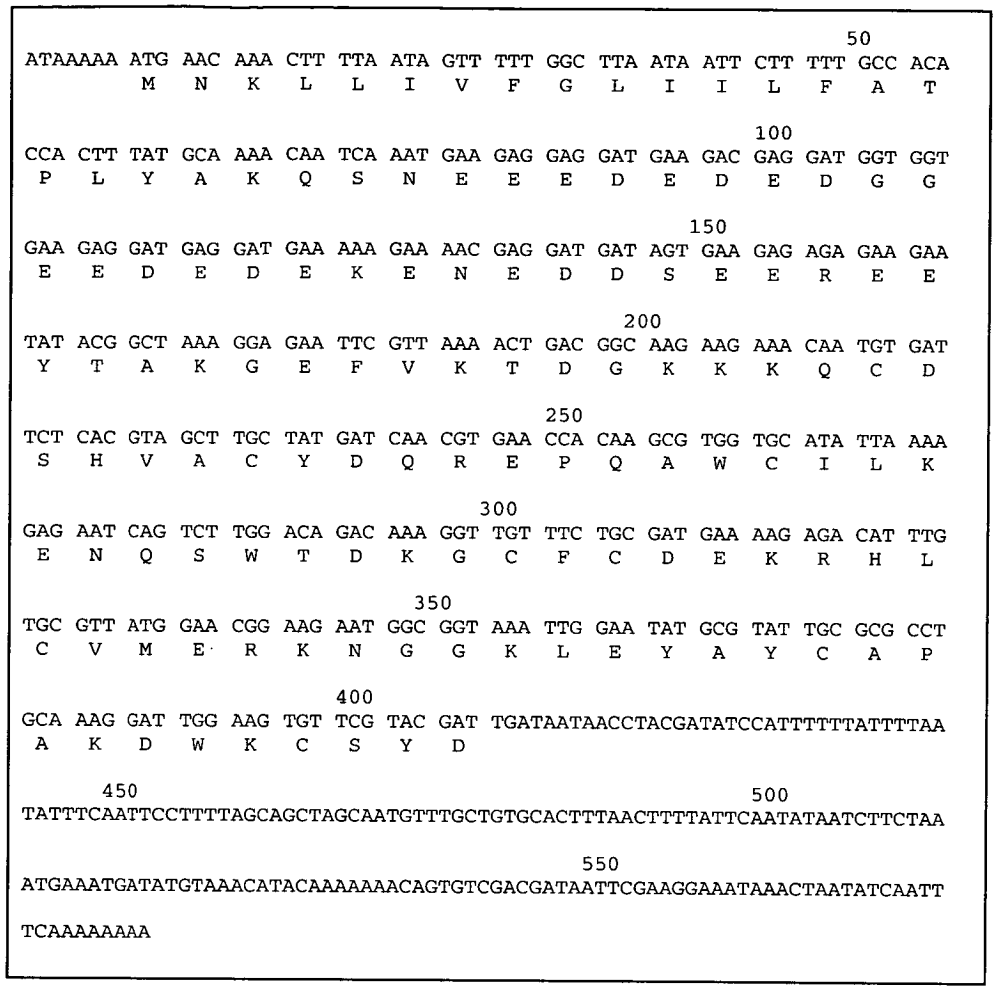


Figure 5.4: Nucleotide sequence and deduced amino acid sequence of Bm-alt-6. *This represents the confirmed sequence of the expressed sequence tag MB3D6AA5G7T3 (GenBank accession number AA841275)*

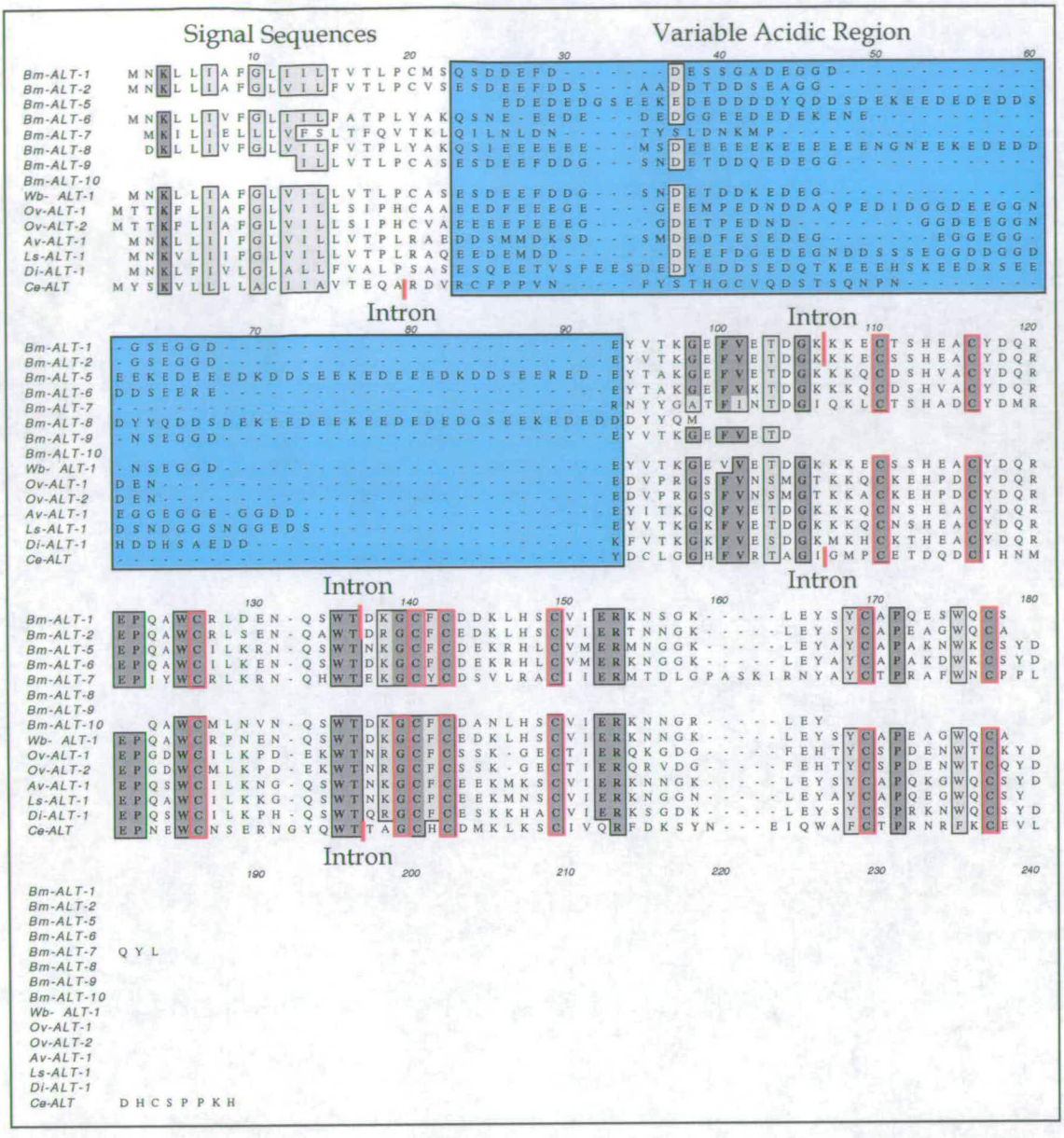


Figure 5.5: Alignment of the predicted protein sequences of the filarial alt genes with the proposed homologue from C. elegans. *Vertical red lines indicate intron position where known. The eight conserved cysteine residues are boxed in red. GenBank accession numbers are a follows: Bm-ALT-1; U57547, Bm-ALT-2; U84723, Wb-ALT-1; AF0845533, Av-ALT-1(Av-18); U47545. Di-ALT-1(Di-20/22); U29459, Ov-ALT-1; U96176, Ov-ALT-2; AF044952 and Bm-ALT-10; U29578. The Ls-ALT-1 sequence is generated from a consensus of 18 Litomosoides sigmodontis ESTs (Allen et al manuscript submitted). The C. elegans sequence is a translation of sequence from cosmid CO8A9 with conformation of the predicted open reading frame by cDNA sequencing. Bm-ALT-5 and -6 are full sequences of clones identified by EST sequencing (SW3D9CAC340 and MB3D6AA5G07 respectively). Bm-ALT-7is the full sequence of EST clone SWAMCAC14H05 (David Guiliano, unpublished). Bm-ALT-8 is derived from consensus sequences from the following EST sequences SW3ICA1521SK, AS3ISB330SL, SW3ICA107SK, SW3ICA1123SK, BSBmL3SZ44N11SK and BSBmL3SZ45B6SK). Bm- ALT-9 is a translation of EST MB3D6AA0H07T3. Note that Bm-ALT-5, -8, -9 are probably not full length sequences as the clones or clusters do not encode potential initiation methionines.*

One clone from a day 9 infective larvae library was chosen for sequencing, SW3D9CA340SK (GenBank accession number AA585623). The gene differs significantly from *alt*s-1 and -2 (46 and 43% amino acid identity respectively) and was therefore designated Bm-*alt*-5 (Figure 5.3). The gene has unusually long C-terminal acidic region (in excess of 68 residues) but lacks both an initiation methionine and a signal sequence, probably due to truncation during preparation of the library. Two additional ESTs (BSBmL3SZ15H18SK and SWYACAL01C115´) appear to derive from this gene along with a partial genomic sequence (GenBank U29577, Wisnewski, N., Frank, G.R. and Grieve, R.B., unpublished). The genomic sequence reveals the conserved position of intron 2 in this and the *alt*-1 and -2 genes. Clone SWYACAL01C115´ comes from a young adult library constructed is unusual as it is one of the few *alt*-family transcripts detected from stages other than L3.

The full sequence of EST clone MB3D6AA5G07 from the day 6 infective larvae library contains a complete open reading frame with 68% identity to the *Bm-alt*-5 protein sequence. The gene, designated *Bm-alt-6* (Figure 5.4), also matched 3 further ESTs (MB3D6VH7H0T3, MB3D6AA7E11T3 and MB3D6V5A10T3) all selected from the day 6 library.

Nine other EST could be identified from a database search. One, from the adult male library did not group with any other ESTs. Full sequencing of the insert (D. Guilliano, unpublished) showed a predicted full length gene. In comparison with the other ALTs the predicted protein from this gene, designated Bm-*alt*-7, lacks an terminal acidic region although the central and C-terminal regions are conserved.

A cluster of 7 ESTs from the infected larvae library defines a gene, designated *Bm-alt*-8. The protein sequence predicted from a consensus of the 7 ESTs consists of approximately 20 amino acids similar to the ALT signal sequences and along with 70 predominantly acidic residues, however the sequences do not extend into a conserved central and C-terminal region.

Two short sequences define two further ALT family members. A single EST from the day 6 library spanning only 152 nucleotides defines *Bm-alt*-9 and a short

genomic sequence (U29578, Wisnewski,N., Frank,G.R. and Grieve,R.B., unpublished) has been designated *Bm-alt*-10.

As the *B. malayi* EST project proceeds further members of the *alt* family will undoubtedly be identified. The genes identified to date can to be broadly divided into 2 groups on the basis of the abundance (as assessed from EST data) and the timing of their expression. The first group, containing *alt*-1 and -2 are highly expressed in the vector derived L3, with lower expression following transmission. The second group are much less abundant, and are largely confined to libraries made from post-parasitic larvae and adult worms

5.2.3 Analysis of *alt*-1 and -2 genomic structures

Primer pairs designed to amplify the *alt*-1 and -2 genes from the start of the predicted mature protein to the 3´UTR were used to amplify genomic DNA. The products were cloned and sequenced which allowed a comparison of intronic structure. Both genes contained 2 introns, which show no similarity between the two genes, with intron/exon boundaries conforming closely to the consensus described for *Brugia* genes (Zang et al., 1999). However, two products were observed from genomic PCR for *alt*-2 (Figure 5.6), and this was found to be due to a dimorphism in the second intron of *alt*-2. Both *alt*-2 introns are unusual in consisting largely of 27 or 46 bp repeat units (Figures 5.9 and 5.10). The dichotomy in *alt*-2 is due to the presence of either 6 or 9 copies of the 46 bp repeat (Figure 5.7).

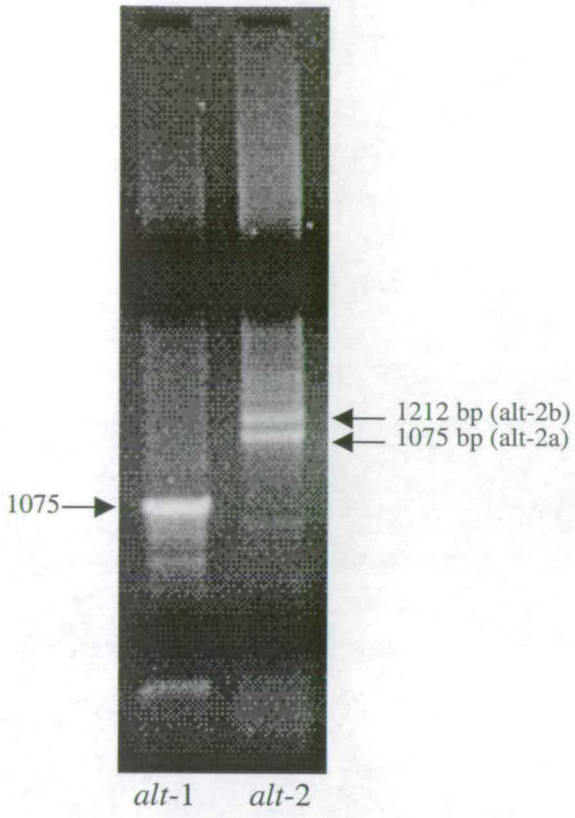


Figure 5.6: Genomic PCR using gene-specific primers . *Genomic DNA from mixed sex adults amplified with lane 1 : alt-1-specific primers (alt-1FOR and alt-13´UTR) and lane 2 alt-2-specific primers (alt-2FOR and alt-2 3´UTR.*

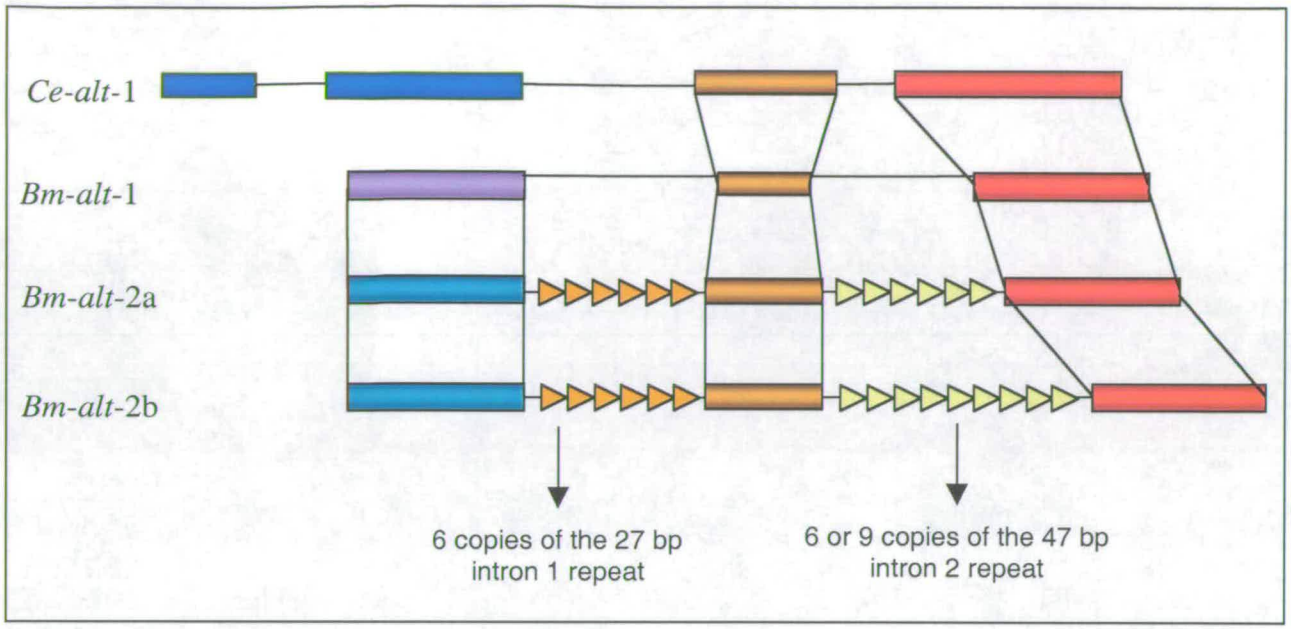


***Figure 5.7: Schematic comparison of the partial Bm-alt and Ce-alt genomic sequences**: Exons with significant homology are coloured identically. Each homologous repeat unit is represented by a identically coloured triangle.*

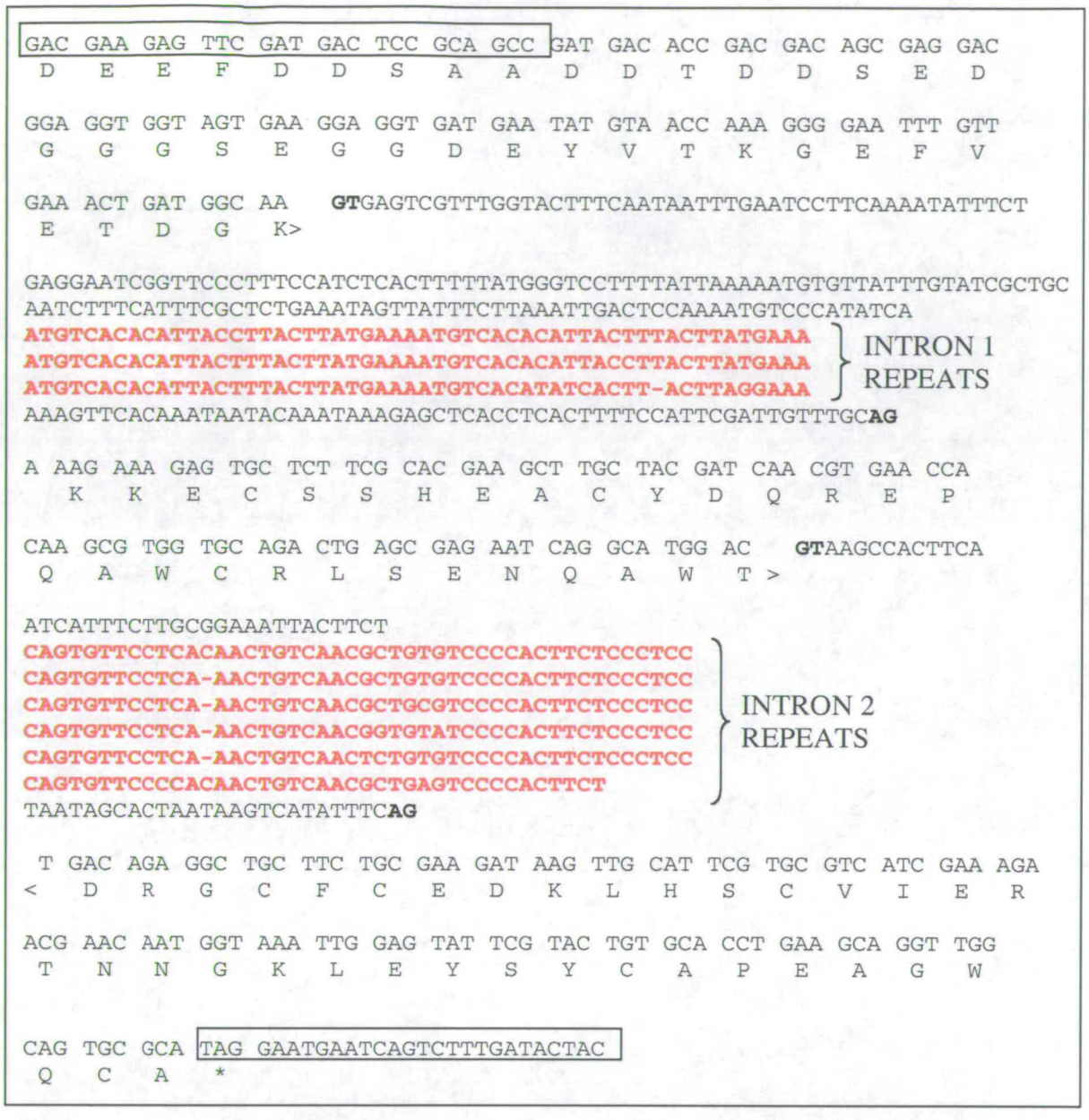


Figure 5.9. Partial genomic sequence of Bm-alt-2a (lower band). *The nucleotide sequence is accompanied by the amino acid translations of the 3 exons. Oligos used to amplify the fragment from genomic DNA are boxed. The repeated sequences within the introns are highlighted in red. Intron splice sites are indicated in bold. > indicates a split codon. This sequence has been deposited with GenBank under the accession number AF183573*

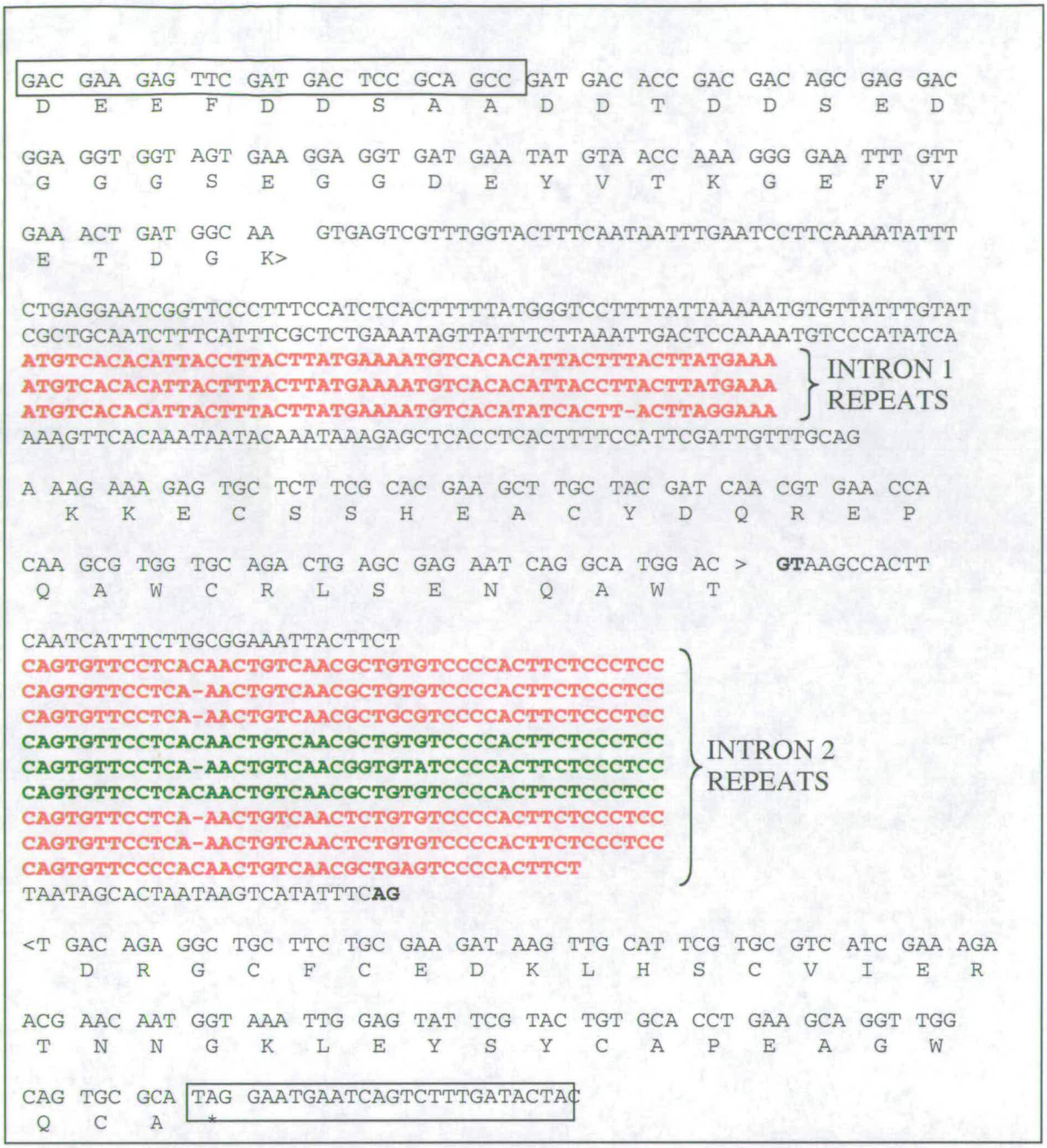


Figure 5.10. Partial genomic sequence of Bm-alt-2b (upper band). *The nucleotide sequence is accompanied by the amino acid translations of the 3 exons. Gaps have been introduced in the repeat intron 2 repeat unit to optimise the alignment. Nucleotides that differ in alt-2a and b are shown in green. Oligos used to amplify the fragment from genomic DNA are boxed. The repeated sequences within the introns are highlighted in red. Intron splice sites are indicated in bold. > indicates a split codon. This sequence has been deposited with GenBank under the accession number AF183574.*

5.2.4 Diversity within the *alt*-2 gene.

The two *alt-2* genomic PCR bands were obtained by amplifying DNA isolated from a large number of mixed sex adult worms. It was unclear whether the two *alt*-2 products were allelic variants of one gene or two distinct genes. Single worm PCRs were carried out to resolve this issue. Genomic DNA products from single adult males revealed a polymorphism within the *alt*-2 gene (Figure 5.11). Half of the single worm DNA produced the 1075 and 1212 bp bands produced by the mixed sex DNA. Three individuals appeared homozygous: 2 contained only the 1212 bp *alt*-2b gene while one contained only the 1075 bp *alt*-2a gene. Additional bands were seen at approximately 1.4 and 1.5 kb in some individuals: one individual contained only this doublet while another produced bands at 1.0, 1.4 and 1.5 kb. A 625 bp band was also present in most individuals. Sequencing of these additional bands is required to positively identify then as allelic variants of *alt*-2.

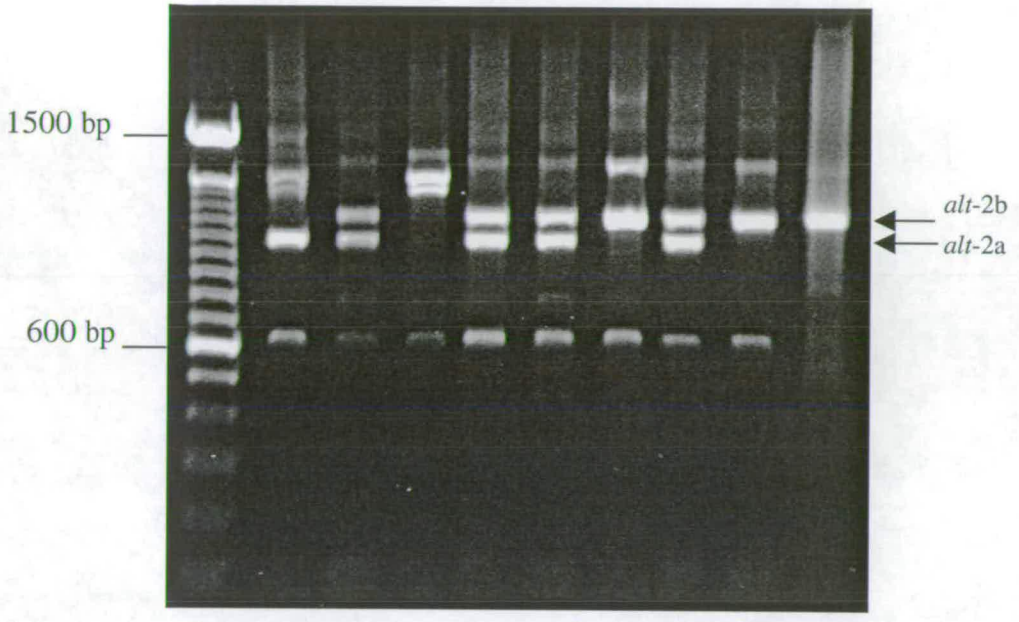


Figure 5.11: Heterogeneity of the alt-2 gene. *Lane 1, 100 bp ladder (Gibco-BRL). Lanes 2-9, single adult male worm genomic DNA amplified with alt-2-specific PCR primers (alt-2 FOR and alt-2 3′UTR). Lane 10, Positive control: amplification of a plasmid containing an alt-2b insert with the same primer pair.*

5.2.5 An *alt*-like gene in *C. elegans*

alt-1 and *alt*-2 cDNA sequences encode novel proteins each with signal sequences and N-terminal acidic tracts, which share 79% amino acid identity (Gregory et al., 1997) and similar proteins have been described in other filarial species (Frank et al., 1996; Joeseph et al., 1998; Pogonoka et al., 1999). Alignment of the filarial ALTs illustrates that the acidic tract is highly variable while the remainder of the protein is conserved (Figure 5.5). Among the predicted protein sequences from *C. elegans* deposited in Wormpep (http://www.sanger.ac.uk/Projects/C_elegans/wormpep), only distant similarities could be found with low significance, including two genes annotated as phospholipase A2. However when the complete nucleic acid sequence of *C. elegans* was searched a more significant similarity ($P>10^{-6}$) was found on cosmid C08A9. This cosmid contains a short tract corresponding to the second and third exons of *Bm-alt*-1 and -2 and encodes a predicted peptide sequence with 32% amino acid identity with *Ce*-ALT-1 (Fig. 5.5). Moreover, C08A9 shows exact alignment of all cysteine residues with ALT-1 and –2 and includes introns in identical positions to those determined for *alt*-1 and –2.

To confirm the predicted intron splice site and to assess whether sequence is transcribed primers designed to amplify across the two *C. elegans* exons to PCR amplify poly A-primed first strand cDNA. Figure 5.13 demonstrates that the gene is transcribed and sequencing of the PCR product confirmed that the splicing predictions are correct. The gene has therefore been designated Ce-*alt*-1.

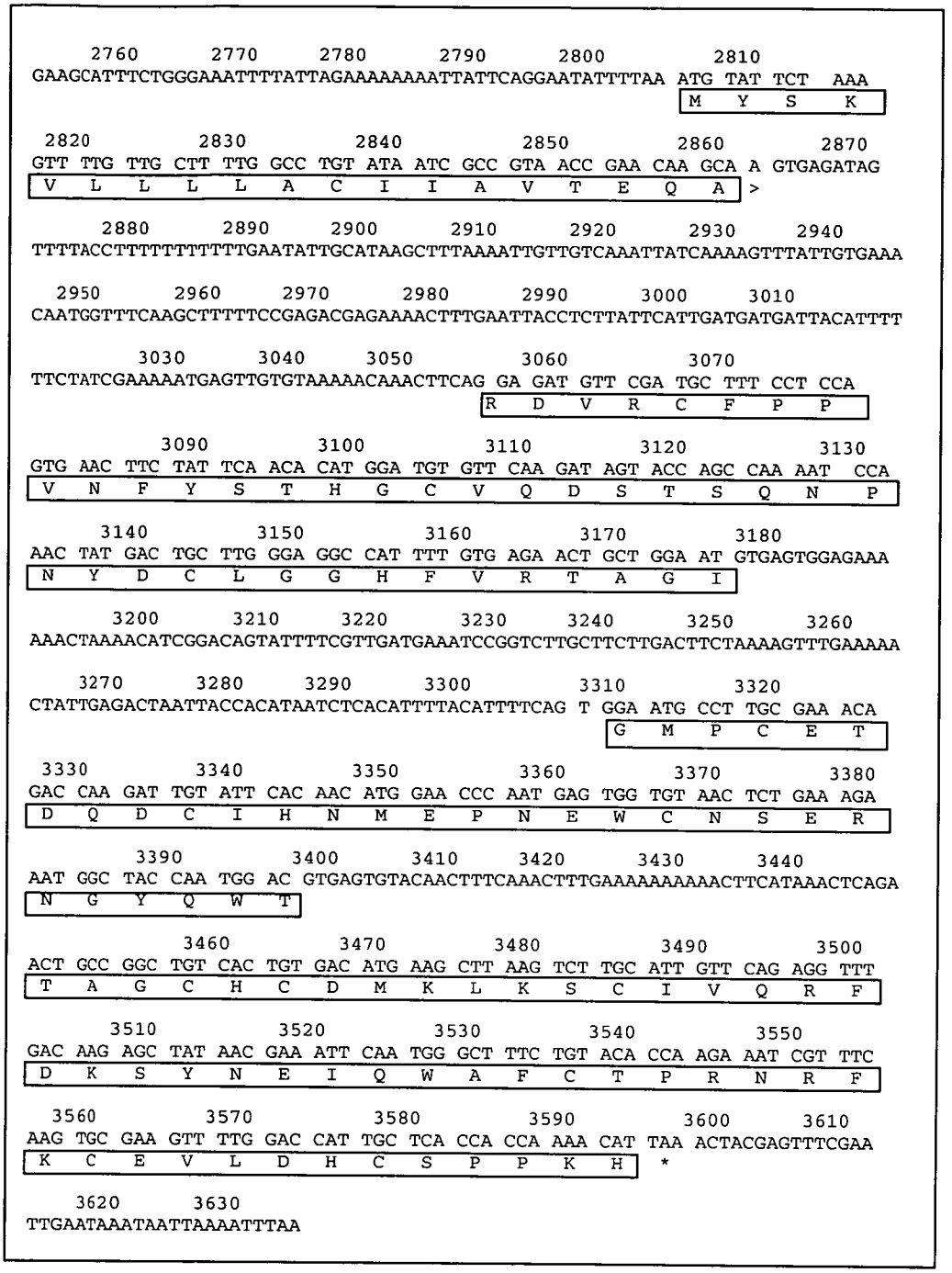


Figure 5.12: Genomic and predicted amino acid sequence of Ce-alt. *The amino acid sequence was deduced by sequencing of cDNA. The sequence starts at base pair 2752 representing the longest transcript identified from 5′ RACE products. The cDNA is polyadenylated at base 3635. Split codons are indicated by >.*

No upstream sequence from the cosmid was similar to the N-terminal acidic domain. The bases immediately 5′ to the first predicted exon correspond to the consensus intron splice acceptor site TTT CAG (Blumenthal and Steward, 1997) suggesting that the identified sequence is incomplete. Amplification of cDNA with a gene-specific reverse primer and a SL-1 primer and failed to produce a product as

did amplification of a mixed stage library using a specific primer in combination with a plasmid vector primer. 5´RACE using the GeneRacer Kit (Invitrogen) identified a cDNA corresponding to nucleotides starting at base 2753 of cosmid C08A9 and lacking a 5´SL sequence. The 5´RACE products allowed the identification of 60 extra N-terminal residues including a signal peptide and an additional intron in the upstream sequence (Figure 5.5). It is not clear whether an intron also exists in the same position in *Bm-alt*-1 and -2. The section of *alt*-1 and -2 amplified to generate genomic sequence lies 6 nucleotides downstream of the intron position in *C. elegans*. Comparison of the N-terminal sequences showed divergence in the amino acid content of the Ce-ALT predicted protein: no stretch of acid residues is present. Bm-ALT-7 is the only other sequence identified to date that lack acidic residues in the N-terminus. A phylogram comparing only the conserved core regions of the filarial and *C. elegans* sequences also demonstrates that Ce-ALT is most close to ALT-7 (Figure 5.14)

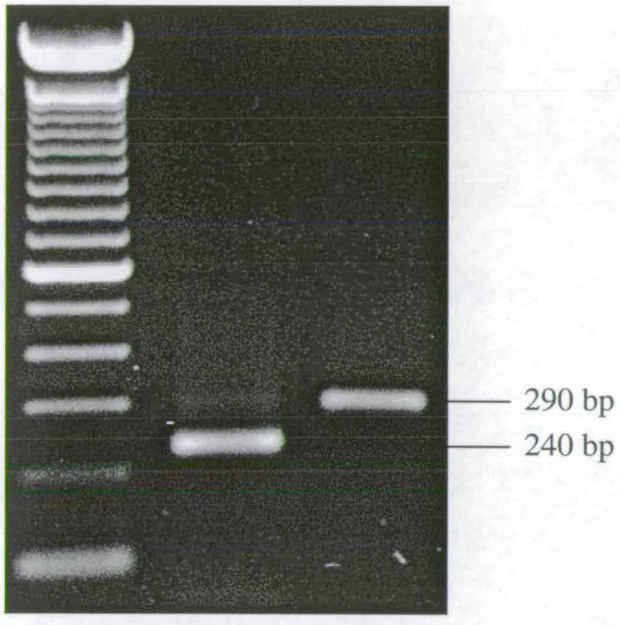


Figure 5.13: Ce-ALT-1 is an expressed gene. *Lane 1: 100 bp ladder (Gibco BRL). Lane 2: A band of 240 bp is generated by PCR of mixed-stage C. elegans poly A-primed first strand cDNA with gene-specific primers Ce-altFOR and Ce-altREV. The primers are designed from sequences in exons 1 and 2 respectively and therefore span the predicted 50 bp intron. Lane 3: PCR of C. elegans genomic DNA using the same primers.*

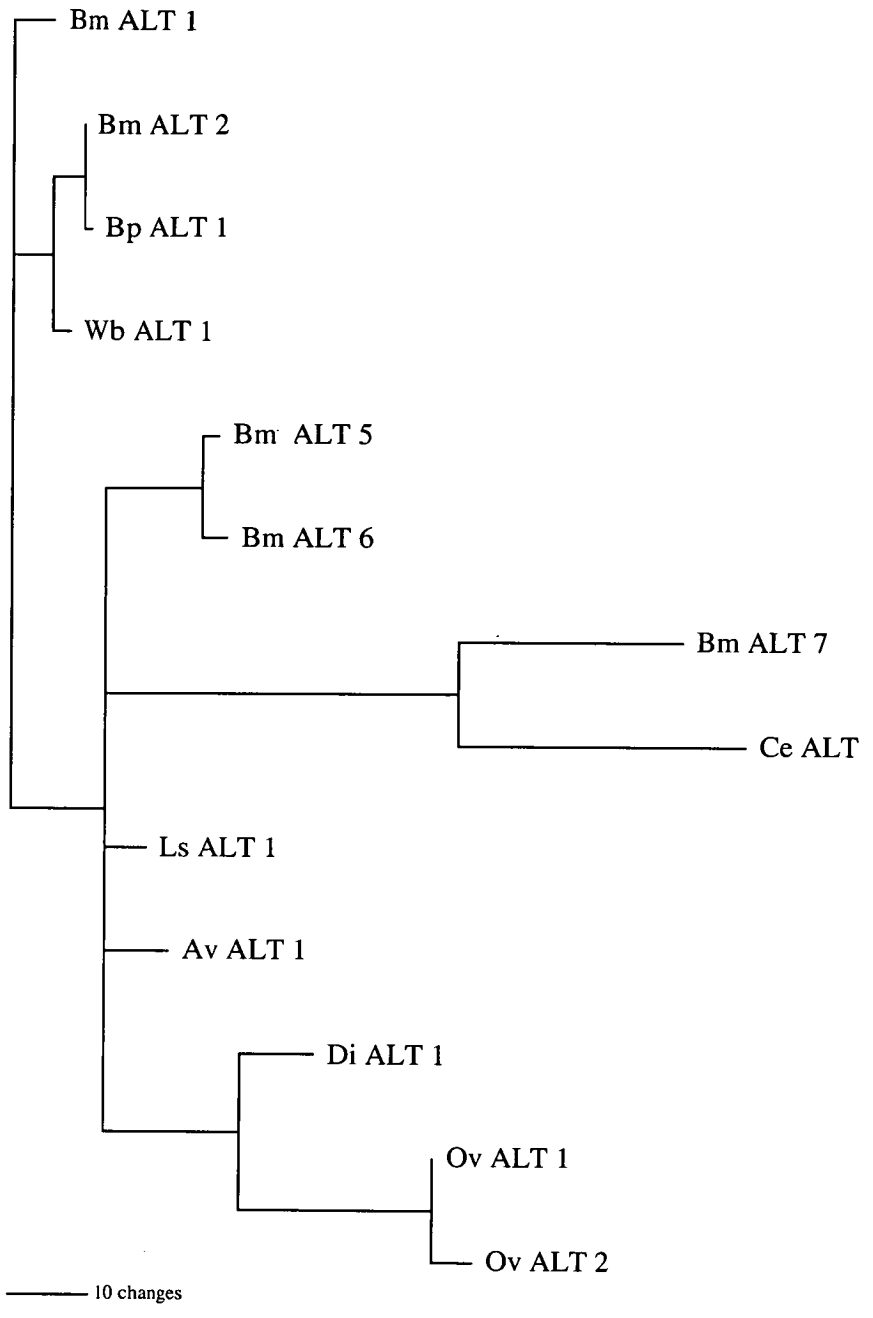


Figure 5.14: Unrooted phylogram showing relatedness of the ALT conserved core in filariae and C. elegans. *The conserved core is defined as the central and C-terminal residues starting at position 100 in the multiple alignment shown in Figure. 5.6.*

5.2.6 Stage-Specific Gene Expression

Expression of *alt*-1 and *alt*-2 at different points of the filarial life cycle was assessed by RT-PCR. For *alt*-1 and *alt*-2-specific PCRs primers binding to the region corresponding to the first 9 amino acids of the mature polypeptides were paired with primers binding to the 3´UTR of the genes. PCR was carried out on poly A-primed first strand cDNA taken at daily intervals during development of parasites from the microfilarial stage to the infective larva in *Ae. aegypti* vector mosquitoes (Figure 5.15). This showed that both *alt*-1 and -2 are switched on between 5 and 6 days following uptake into the mosquito vector, and remain expressed for the duration of tenure in the insect. Similarly, parasites were recovered following infection of the rodent host *M. unguiculatus*. Here, *alt*-1 expression terminated abruptly on transfer into the jird, although brief periods of transcription were detected between days 4 and 8, no subsequent expression could be detected (Figure 5.16). *alt*-2 transcription was less rigorously controlled, with expression continuing for 3 days post infection, and recurring at intervals over the following 3 weeks.

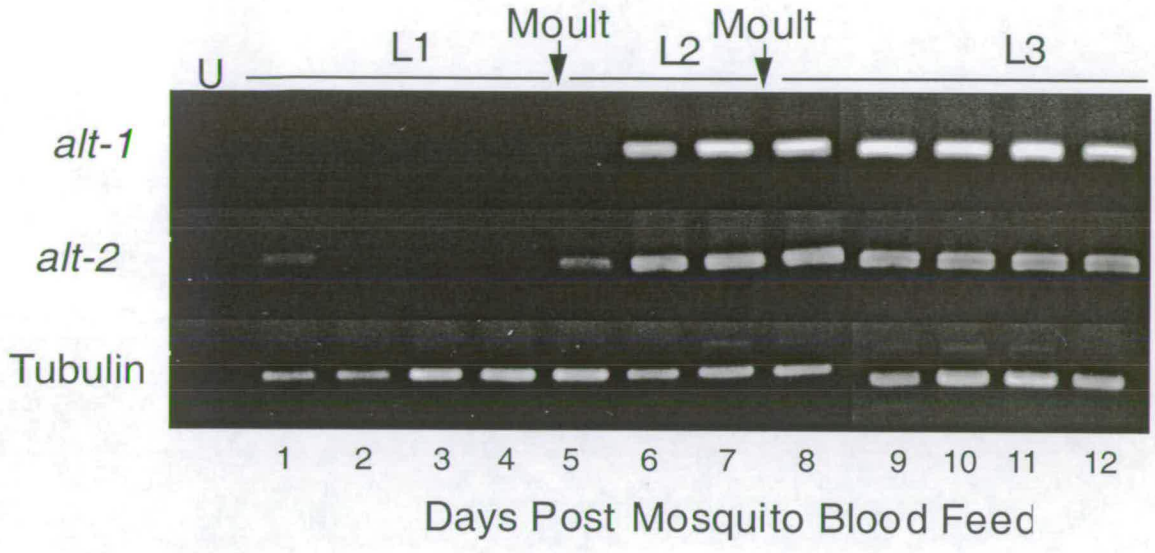


Figure 5.15: ***RT-PCR analysis of the expression of alt-1 and -2 and during larval development within the mosquito vector.*** *U represents RNA extracted from mosquitoes fed on uninfected blood.*

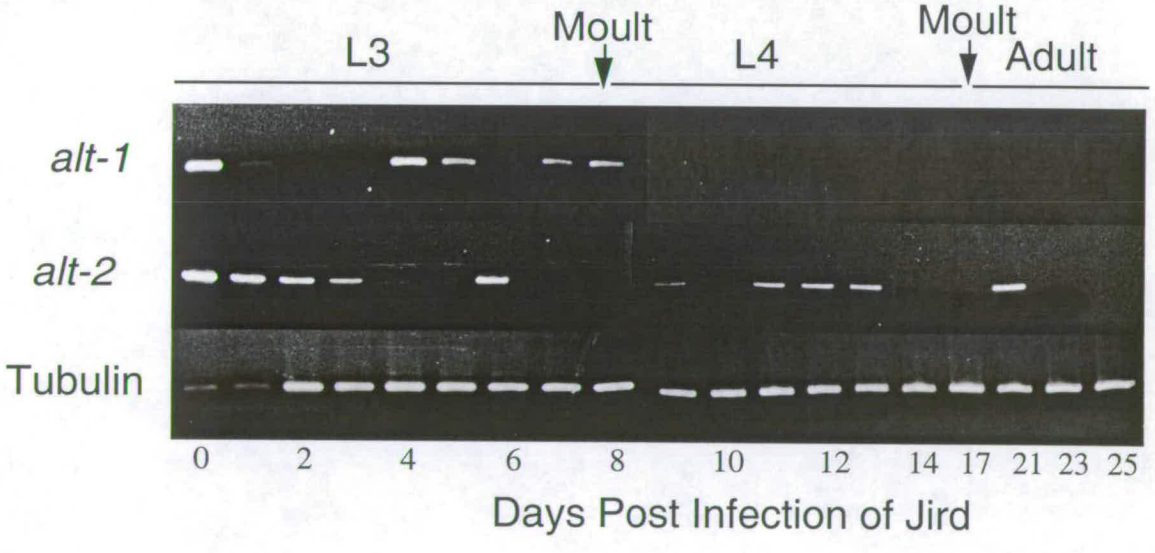


Figure 5.16: ***RT-PCR analysis of the expression of alt-1 and -2 during the first 25 days of infection of the gerbil.*** *Day 0 represents RNA extracted from vector-derived larvae.*

5.2.7 Expression of recombinant ALT.

The predicted full length mature ALT-1 protein was expressed using the pET-29T system (Novagen). Primers altpETFOR and altpETREV were used to amplify the gene for the L3 library (SAW94WL-BmL3). This product was then ligated into the T-ended expression vector pET-29T which produces fusion proteins with an N-terminal S•Tag and a C-terminal hexahistidine Tag. Expression in BL-21 DE3 cells resulted in the production of soluble protein which was purified via the fusion protein's histidine tag on a metal chelating column under non-denaturing conditions. The purified protein ran at a molecular weight of approximately 25 kDa as judged by SDS-PAGE, higher that the predicted size of 17.4 kDa (Figure 5.17).

Antibodies to recombinant ALT-1 protein reacted specifically with a doublet of 20 kDa in soluble extract of L3 on Western blots (Figure 5.18), but no reactivity was detectable towards extracts of Mf and adult stages. Thus, at the protein level, ALT-1/-2 are effectively L3-specific. Thus ALT-1 and –2 larval-specific expression parallels that of related proteins in *D. immitis* (Frank and Grieve, 1991; Frank and Grieve, 1996; Frank et al., 1996) and one of the *O. volvulus* homologues, ALT-1 (Joeseph et al., 1998). However the expression contrasts with alt family members from *A. viteae* (Pogonoka et al., 1999) and Ov-ALT-2 (Joeseph et al., 1998).

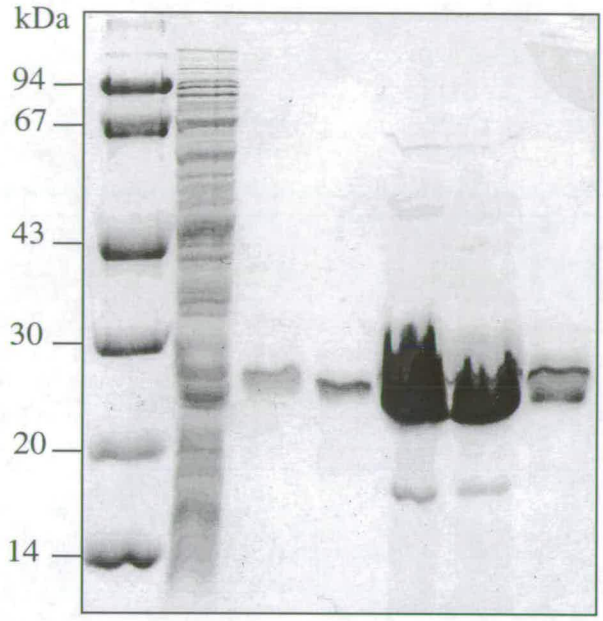


Figure 5.17. Expression and purification of ALT-1. *Lane 1: Molecular weight markers (Pharmacia). Lane 2: total extract of induced cells. Lanes 3-7: sequential fractions eluted from the His•Bind purification column.*

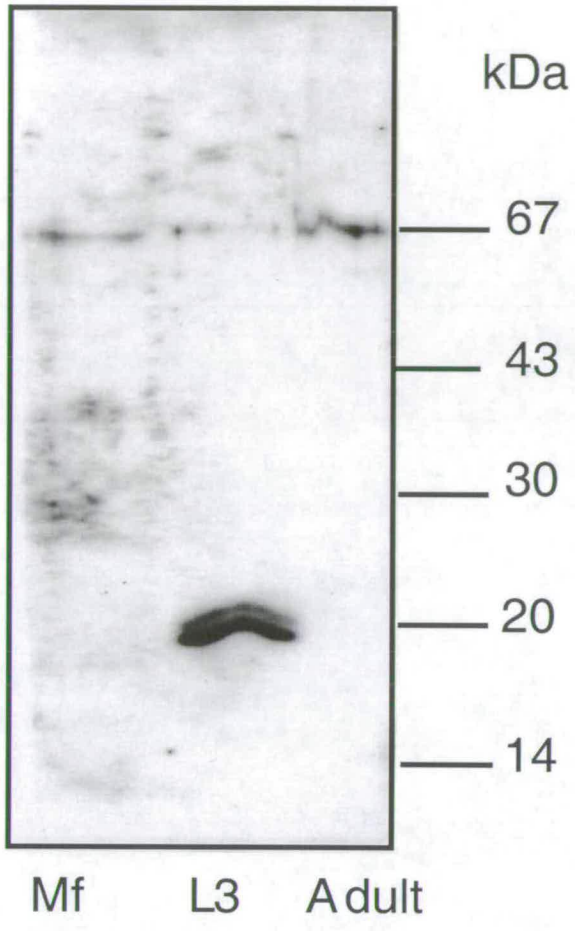


Figure 5.18: ***Detection of ALT-1 in extracts from different stages of the life cycle.*** *Detergent-soluble extracts of microfilariae, infective larvae and adult B. malayi were separated on a 15% SDS-PAGE, transferred to Immobilon™ (Millipore) and probed with antisera raised against recombinant ALT-1.*

5.2.8 Human Recognition

The prominence of the *alt* transcripts suggests that exposed humans may be serologically reactive to the ALT proteins. This was tested with sera from 40 patients resident in a *B. malayi*-endemic area of Indonesia, drawn equally from the two categories of amicrofilaremic and microfilaremic. The former group will contain both parasite-free individuals and patients with subpatent infections; the latter group all have detectable blood-borne microfilariae. Because anti-filarial antibodies, measured to adult and microfilarial-stage antigens, are dominated by the IgG4 isotype (Hussain et al., 1987; Kurniawan et al., 1993; Maizels et al., 1995), all four IgG isotypes were individually assayed.

Antibodies to rALT-1 protein are found in members of both groups (Figure 5.19). Interestingly, the antibody isotypes are predominantly IgG1 and IgG3, and no IgG4 is observed. This contrasts with the profile observed for a *B. malayi* protein expressed by the adult stage, Bm33 (Dissanayake et al., 1993), which is the target of IgG4 isotype antibodies. There is also a discordance with observations on the *O. volvulus* protein Ov-ALT-2 (86% identical to but 8 amino acids smaller than Ov-ALT-1). This is constitutively expressed (Joeseph et al., 1998), and human onchocerciasis patients have high levels (95%) of seropositivity, including IgG4. Thus, in *O. volvulus* ALT does not represent a larval-specific antigen.

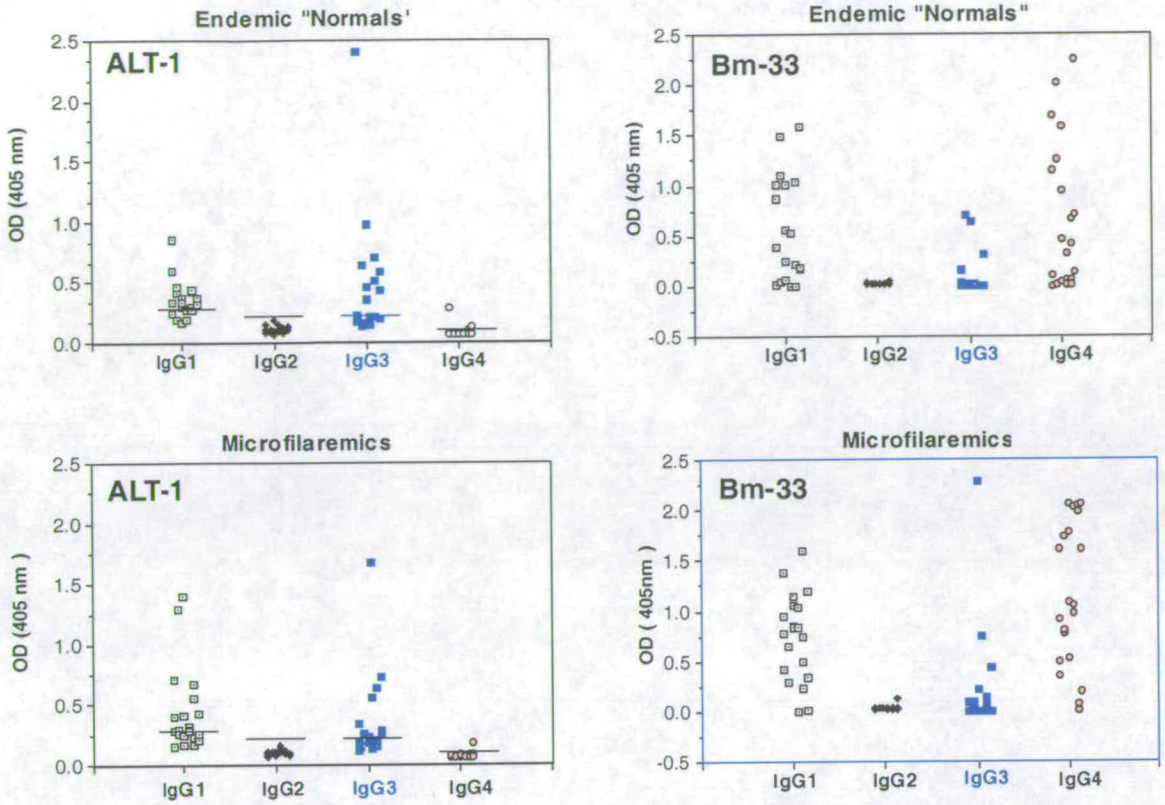


Figure 5.19: Bm-ALT-1 is not the target of IgG4 antibodies during human infection. *Elisa using 40 human B. malayi infection sera against Bm-ALT-1 and Bm-33 recombinant proteins. Upper panels show sera from amicrofilaremic endemic normals; lower panels show sera from patients with circulating microfilariae. Horizontal bars in the ALT-1 panels represent the mean value +3 standard deviations of the non-endemic human control sera. Values for Bm-33 represent the net OD value following the subtraction of readings for each serum against the maltose binding protein fusion partner. ELISA assays and preparation of Bm-33 recombinant were performed by Dr A. Kurniawan-Atmadja*

5.2.9 Protective Immunisation

To assess the potential of ALT-1 for vaccination, gerbils were immunised jirds with 4 doses of ALT-1 spread over 10 months, and challenged them with 300 live L3. Four weeks later, the recovery of live parasites in the immunised group was 76% reduced compared to the adjuvant-alone controls (Figure 20). This difference was significant at the $p<0.05$ level by Whitney-Mann nonparametric statistics. The data reported here indicate greater than 70% protection in jirds against a challenge infection; this is substantially better than any previous filarial recombinant antigen reported, and is in the range achieved by vaccination with radiation-attenuated larvae, of 44-91% (Yates and Higashi, 1985). Of previously tested antigens, paramyosin has yielded disappointing results (Li et al., 1993; Li et al., 1999), while heat shock protein 70, myosin and α1-type IV collagen have recently been shown not to stimulate protective immunity (Peralta et al., 1999). Thus, ALT-1 and ALT-2 offer the best vaccine candidates yet found for filariasis.

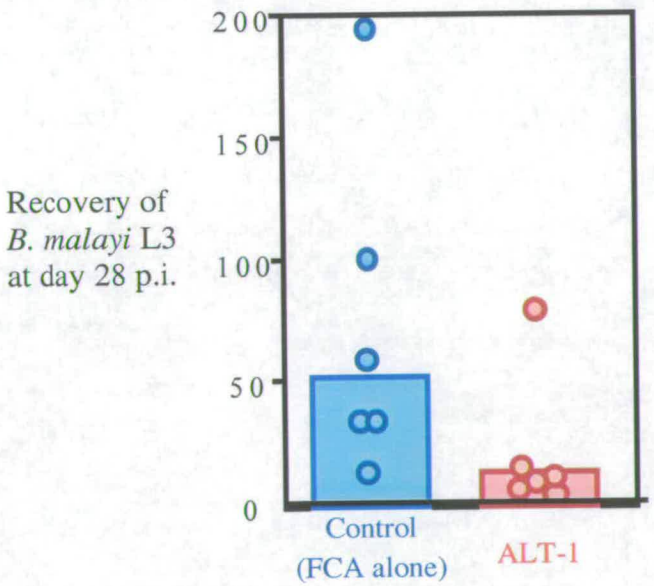


Figure 5.20. Bm-ALT-1 reduces recovery of parasites from vaccinated animals by 76%. *Challenge parasites were recovered 28 days after an intraperitoneal infection with 300 B. malayi infective larvae. Circles represent recoveries from individual animals. The red and blue boxes represent the mean recovery from each group.*

DISCUSSION

Vaccination against helminth parasite organisms has proved problematic, both in identifying likely vaccine antigens from the wide repertoire of antigens expressed, and with respect to the immunopathological responses to many of these antigens (Maizels et al., 1999). A new approach is reported here, of selecting highly-expressed, stage-specific products such as the ALT proteins, which are shown not to be present in the mature adult stage. This, coupled with the fact that ALT proteins are parasite-specific products unrelated to any host constituent, renders less likely any adverse consequences of immunisation. There is evidence from two other filarial species to associate ALT recognition with immunity (Frank and Grieve, 1991; Frank and Grieve, 1996; Frank et al., 1996; Joeseph et al., 1998), but the data provided here provide the first demonstration of protective immunity in a susceptible host using a single recombinant protein. The high level of sequence similarity between ALT sequences from *B. malayi* and *W. bancrofti* suggest that there will be immunological cross-protection between these two species, of which *W. bancrofti* is responsible for >90% of human infections but which does not infect laboratory animals.

The *C. elegans* gene has not been identified as a predicted gene by GeneFinder because of its small size, the program excludes small open reading frames. In addition the gene is also not been identified within the 75,000 clones sequence tagged from a *C. elegans* libraries. This may be again due to its small size (clones with inserts less than 500 bp are not selected for sequencing) or perhaps the stage specificity of the gene makes it a rare transcript amongst the mixed staged population of nematode mRNA used in the library construction.

The absence of a sequence encoding acidic residues upstream of the Ce-*alt* gene may indicate that this region is specific to parasites and may represent part of their adaptation to parasitism. Alternatively we may have to search further upstream for additional exon. Although *C. elegans* introns tend to be small (more than half are shorter than 60 bp) some very large introns are present, for example the first intron of the *unc*-7 gene is 18 kb (Starich et al., 1993).

The variation seen in the Bm-*alt*-2 gene from single male worms is suggestive of allelic variation although sequencing of the amplified products to confirm that they are *alt*-2 gene products is necessary. This would also confirm whether the size variation is due to diversity in the intron 2 repeats. Repeat units are a common occurrence in the genomes of multicellular organisms. In *C. elegans* tandem repeats account for 2.7% of the genome and approximately half of these sequences are found within introns (The *C. elegans* Sequencing Consortium, 1998). In general, repetitive sequences are found in intragenic regions and are thought to be involved in maintenance and movement of chromosomes. However, within introns repetitive sequences have been linked with control of transcription. Intron 6 of the human interleukin-1 alpha gene contains a 46 bp repeat region capable of binding the transcriptional factor SP1 (Haugen et al., 1989). Significantly 6 alleles have been described containing between 5 and 18 repeats and gene expression correlates negatively with the number of repeats (Bailly et al., 1996).

The human antibody profile supports, in the context of natural exposure, the concept that ALT-1 is an L3-specific antigen as it has previously been shown that responses to the L3 overall are less dominated by IgG4 (Kurniawan-Atmadja et al., 1998). It is notable that both Mf-negative and -positive groups contain individuals seropositive for ALT-1 antibodies. This finding is in keeping with the notion of age-acquired concomitant immunity, in which exposed individuals gain protection against new infection while being unable to eradicate the resident parasites (Day, 1991; Day et al., 1991a; Day et al., 1991b). Thus, both endemic 'normals' and microfilaria-carriers may express protective antibodies, effectively negating the search for protective antigens by comparing antibodies from "immune" and "susceptible" individuals; such a comparison would have excluded ALT-1 from consideration.

The conservation of *alt* genes in all the filarial nematodes so far studied, and the very weak similarity to a single *C. elegans* locus, implies that ALT products are critical in a filarial-specific role. Moreover, this role evidently requires a remarkably high degree of expression at the point of initial entry by parasites into the

mammalian host. If this role is essential to parasite survival, neutralisation by the immune response may be sufficient to ensure protection. Elucidation of the function of the ALT proteins, and the nature of the response induced by vaccination with these antigens, should greatly enhance our understanding of the immunology of filarial infection, and may lead to successful strategies for the control of filarial disease.

CHAPTER 6

The abundant glycine/tyrosine-rich (*agt*) family of genes

6.1 Introduction.

6.2 Results

- 6.2.1 Abundance of *agt-1* mRNA throughout the life cycle
- 6.2.2 Other *B. malayi agt* sequences identified in dbEST
- 6.2.3 A *agt*-like gene family in *C. elegans*
- 6.2.4 Stage-specific gene expression
- 6.2.5 Expression of recombinant AGT

6.3 Discussion

INTRODUCTION

A multilayered extracellular cuticle surrounds all nematodes. Along with the surface coat, it forms a protective barrier between the animal and its environment and plays important roles in maintenance of morphology, motility and (Bird and Bird, 1991; Blaxter and Robertson, 1998). In addition the cuticle of parasitic nematodes plays a critical role in host-parasite relationships as it represents the main body of tissue in contact with the host immune system.

The entire cuticle is sloughed off and replaced at each of the four postembryonic moults giving rise to cuticles that are different in both observable structures and specific protein composition. Cuticle proteins are synthesised by the underlying syncytial hypodermis and once exported to the cuticle they are extensively cross-linked, conferring both elasticity and rigidity to the structure. The proteins can be broadly divided into those that are solubilised by reducing reagents, such as 2-mercaptoethanol, and those that remain insoluble. A large family of collagens are solubilised by reducing agents breaking intrachain disulphide bonds. Initially synthesised as low molecular weight proteins (pre-procollagens) they are further cross-linked by non-reducible bonds to give apparent molecular weights of 50-160 kDa. The non-reducible cross-links are of two types: (1) di- and tri-tyrosine and (2) ε-(γ-glutamyl) lysine. Di-tyrosine cross-links are found within collagen subunits and also in additional cuticle components, termed cuticlin, found in the insoluble fraction of the cuticle. Two genes encoding components of the cuticlin residue, CUT-1 and CUT-2 from *C. elegans* have been described: both contain a high proportion of glycine and tyrosine residues. Recombinant CUT-2 can be cross-linked *in vitro* to high molecular weight product in the presence of hydrogen peroxide and horseradish peroxidase demonstrating a capacity to form. Both proteins have characteristics of highly cross-linked proteins within the insoluble residues of the larval cuticle and eggshell layers of insects.

One abundant transcript identified from infective-stage larvae belongs to a family of small molecules rich in glycine and tyrosine residues from filariae and *C.*

elegans (Chapter 1 and (Gregory et al., 1997)). The transcript, named *Bm-agt*-1 for **a**bundant **g**lycine/**t**yrosine rich protein, predicts a small protein of approximately 6 kDa with an N-terminal signal sequence. Although not homologous to the CUT proteins the high content of glycine and tyrosine and the presence of a signal peptide suggest that they are substrates for cross-linking enzymes in the extracellular spaces of the nematodes, possibly in the cuticle.

6.2.1 A family of *agt*-like proteins in filariae

The *Bm-agt*-1, formerly called *Bm-alt*-3, has been renamed to avoid confusion with the growing family of genes with homology to *alt*-1/-2. The cDNA was first identified as an abundant band amongst vector derived L3 transcripts (chapter 1 and Gregory et al, 1997). The predicted mature 3.8 kDa protein contains 22.2% glycine and 8.3% tyrosine (Figure 6.1). Although not particularly tyrosine rich, a homologue identified amongst *B. malayi* ESTs, originally called *Bm-alt*-4 and now renamed *Bm-agt*-2, is a 5.3 kDa protein which, without its predicted signal peptide, contains 43.1% glycine and 19.6% tyrosine. The two proteins are 34% identical (49% similar). All ESTs representing *agt*-1 come from vector-derived L3 libraries, however *agt*-2 appears less stage-specific with ESTs from both vector-derived L3 and adult male libraries.

The *agt*-1 gene in *B. malayi* is not amongst the most abundant in the EST dataset: 13 ESTs are identified by BLAST searching. Significantly 9 of these sequences are from a library constructed from vector derived L3 SL-1/oligo-dT PCR products (JHU93SL-BmL3) representing 3% of sequences from this library. In contrast only 0.14% of EST from the conventional vL3 library are *agt*-1. This over-representation in the library and amongst the abundant transcripts (chapter 1) may be due to the preferential amplification of small transcripts by taq polymerases or the abundance of this transcript amongst those that receive the SL-1 sequence.

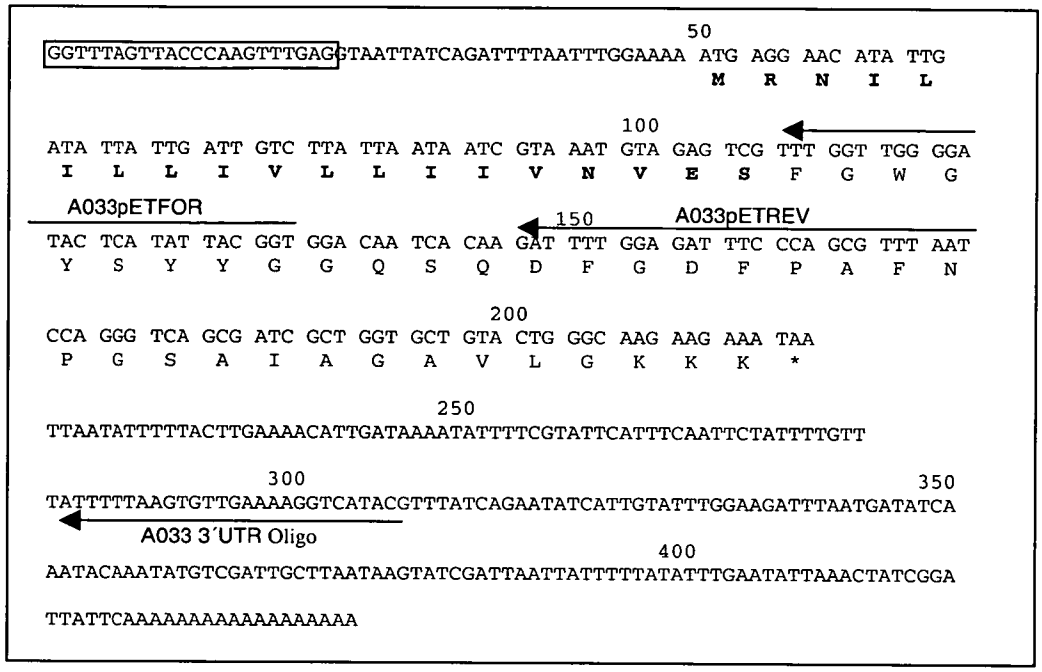


Figure 6.1: Nucleotide and deduced amino acid sequence of Bm-agt-1. *The nematode spliced leader SL-1 is boxed. The putative signal sequence is in bold. Oligo pairs used for RT-PCR (A033pETFOR and A033 3'UTR) and cloning into pET 29 (A033pETFOR and A033pETREV) are indicated.*

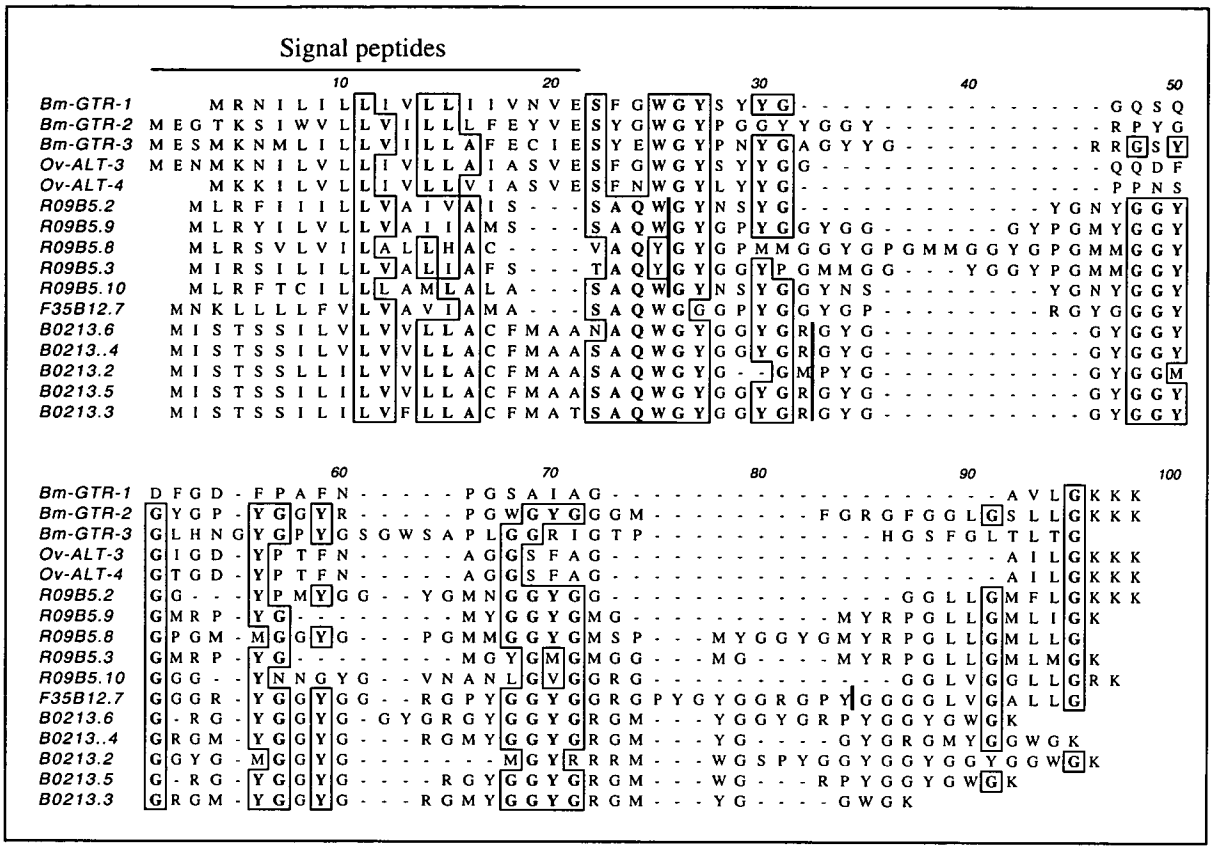


Figure 6.2: Alignment of the five filarial abundant glycine/tyrosine-rich proteins with homologues from C. elegans. *Accession numbers are as follows: Bm-AGT-1; U80974, Bm-AGT-2; consensus of EST sequences R47601, N43123 and N99335, Bm-AGT-3; consensus of EST sequences AI083376, AI92083, AI87702 and AI919654, Ov-ALT-3; Joseph and Lustigman, unpublished, AF020587, OV-ALT-4; consensus of EST sequences AA610980, AA680592, AA294475, AA294034 and AI692149. C. elegans sequences from cosmids R09B5, B0213 and F35B12 are found in GenBank under accession numbers AF039046, AF039050 and Z74032 respectively. Vertical red bars indicate the positions of introns within the C. elegans sequences.*

The closest database hit to *Bm-agt*-1 is an *O. volvulus* gene (Joeseph,G.T. and Lustigman, S., unpublished, GenBank accession number AF020587) with 27% glycine and 10.8% tyrosine residues. In contrast to the *B. malayi* genes the homologue in *O. volvulus* is represented 29 times in the EST dataset: all but one of these sequences are from the vector-derived L3 (vL3) library (0.4% of all ESTs from this library) while only one sequence comes from a library of larvae undergoing the L3-L4 moult. Interestingly a variation of this sequence is also found in the dataset, which we have termed *Ov-alt*-4. Amongst these ESTs the predicted protein varies consistently at nine residues and also includes an insertion of a glycine residue, all in the N-terminal portion of the protein. This variant is much more prominent in the moulting library: 41 ESTs (1.6%) come from this library while only 3 come from the L3 library. The sequence of the variant also lacks the first three N-terminal residues of *Ov*-ALT-3. In *Ov*-ALT-4 the N-terminus starts with MKK (Figure 6.2).

6.2.2 A *agt*-like family in *C. elegans*

The protein sequences for both *Bm*-AGT-1 and -2 were used to identify possible homologues in *C. elegans*. A family of 11 predicted genes was identified. All genes were located on chromosome 5 and clustered into 3 regions covered by 3 cosmids: B0213, R09B5 and F35B12 (Fig. 6.3). Each gene is interrupted by a single intron. Five genes are arranged tandemly on both B0213 and R09B5 although the direction of transcription is distinct: the genes on B0213 are transcribed in the same direction but the middle 3 genes in the R09B5 cluster are in the opposite orientation compared with the flanking genes. This appears to have arisen from an inversion of the middle genes. Both arrays are found downstream of

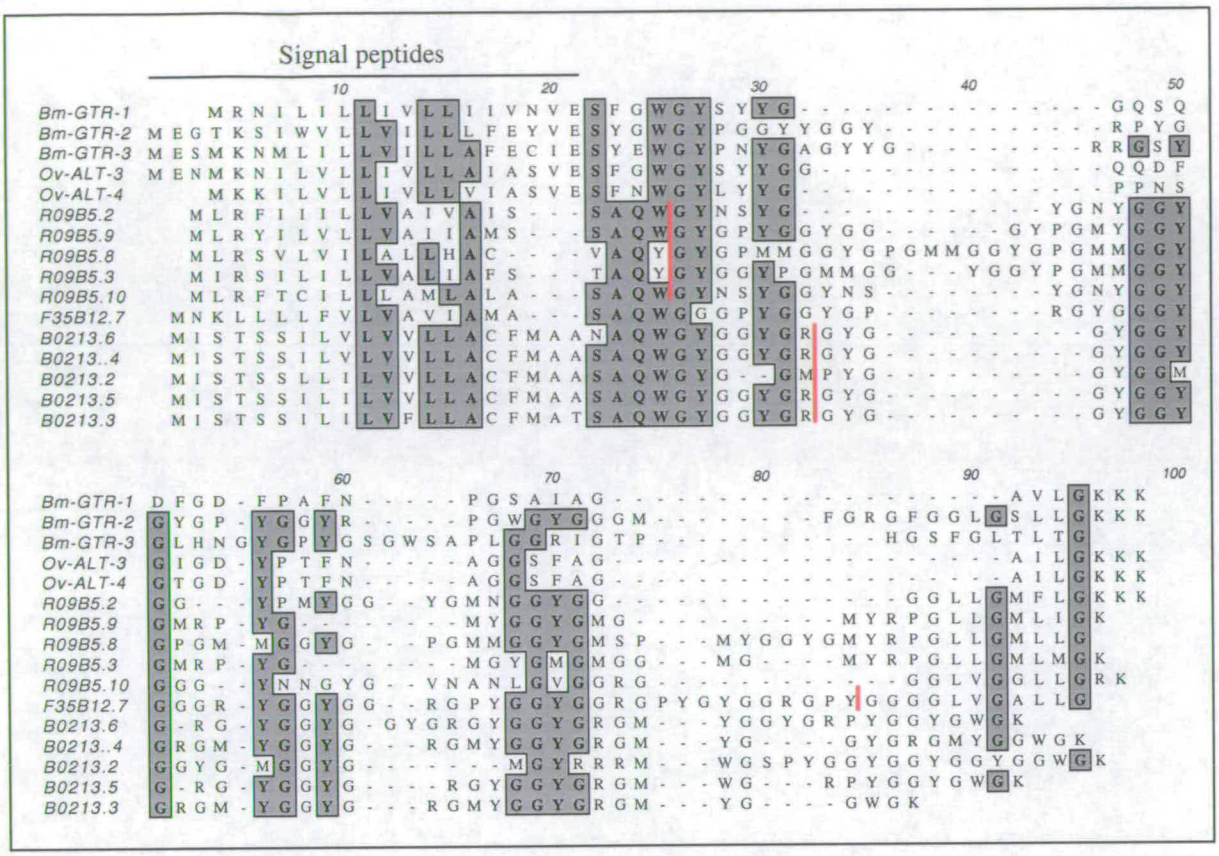


Figure 6.2: Alignment of the five filarial abundant glycine/tyrosine-rich proteins with homologues from C. elegans. *Accession numbers are as follows: Bm-AGT-1; U80974, Bm-AGT-2; consensus of EST sequences R47601, N43123 and N99335, Bm-AGT-3; consensus of EST sequences AI083376, AI92083, AI87702 and AI919654, Ov-ALT-3; Joseph and Lustigman, unpublished, AF020587, OV-ALT-4; consensus of EST sequences AA610980, AA680592, AA294475, AA294034 and AI692149. C. elegans sequences from cosmids R09B5, B0213 and F35B12 are found in GenBank under accession numbers AF039046, AF039050 and Z74032 respectively. Vertical red bars indicate the positions of introns within the C. elegans sequences.*

genes with homology to ODR-10, part of a large family of odorant receptors, suggesting that the two clusters arose from a duplication event and that this synteny has been maintained, at least in part. Also upstream of the genes on cosmid B0213 are six tandem copies of a gene with homology to cytochrome P-450 and an enzyme capable of catalysing the oxidation of tyrosine in the formation of dityrosine. A protein involved in yeast cell wall maturation has significant homology to cytochrome P-450 and is linked to the formation of dityrosine bonds found within the spore wall (Briza et al., 1990).

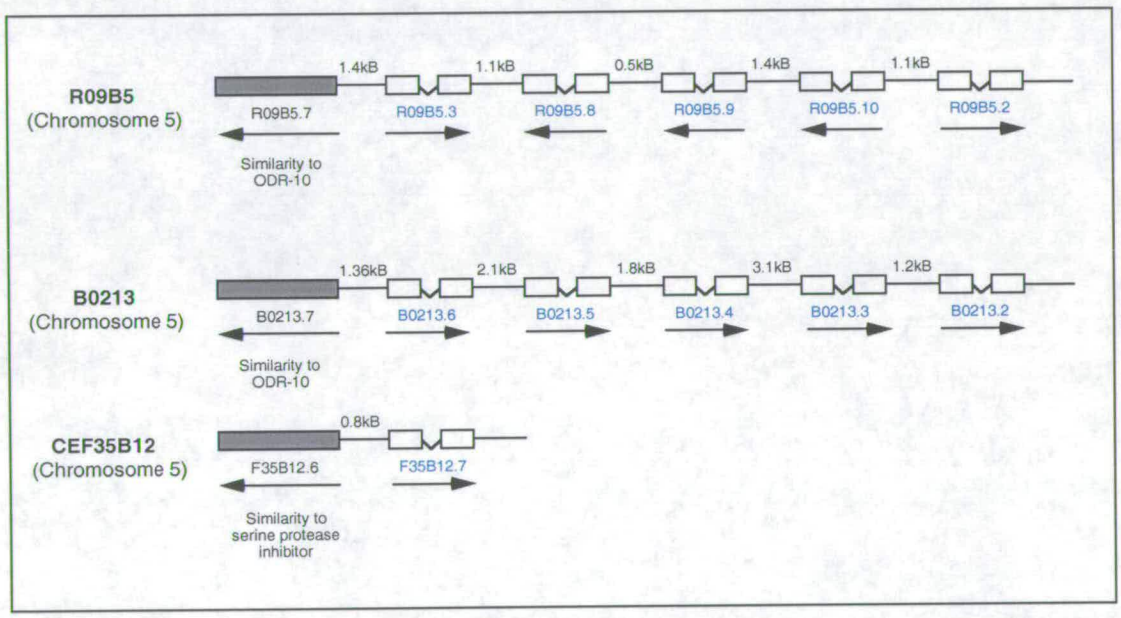


Figure 6.3 : Schematic representation of the organisation of genes on the cosmids containing the C. elegans homologues of Bm agt-1 and -2. *Genes are identified by their cosmid number. Those in blue are agt homologues. Arrows indicate the direction of transcription. Each gene consists of 2 exons and sizes of intragenic regions are also given.*

6.2.3 Analysis of *agt* expression during development using RT-PCR

Expression of *Bm-agt*-1 was assessed by RT-PCR using staged sets of RNAs representing the larval development within the mosquito and also during the first 25 days of development in the mammalian host. This confirmed that expression is largely confined to the L3 stage. Expression begins at day 5 post-infective blood feed and continues until the larvae enter the mammalian host. Expression falls below the sensitivity of the assay within 24 hours of infection. In agreement with the EST data the *agt*-1 signal in vL3's does not appear as abundant as *Bm-alt*-1 or *Bm-cpi*-1 under identical conditions. Larger amplification products are seen at days 0 and 5 of development in the mosquito presumably representing the genomic product.

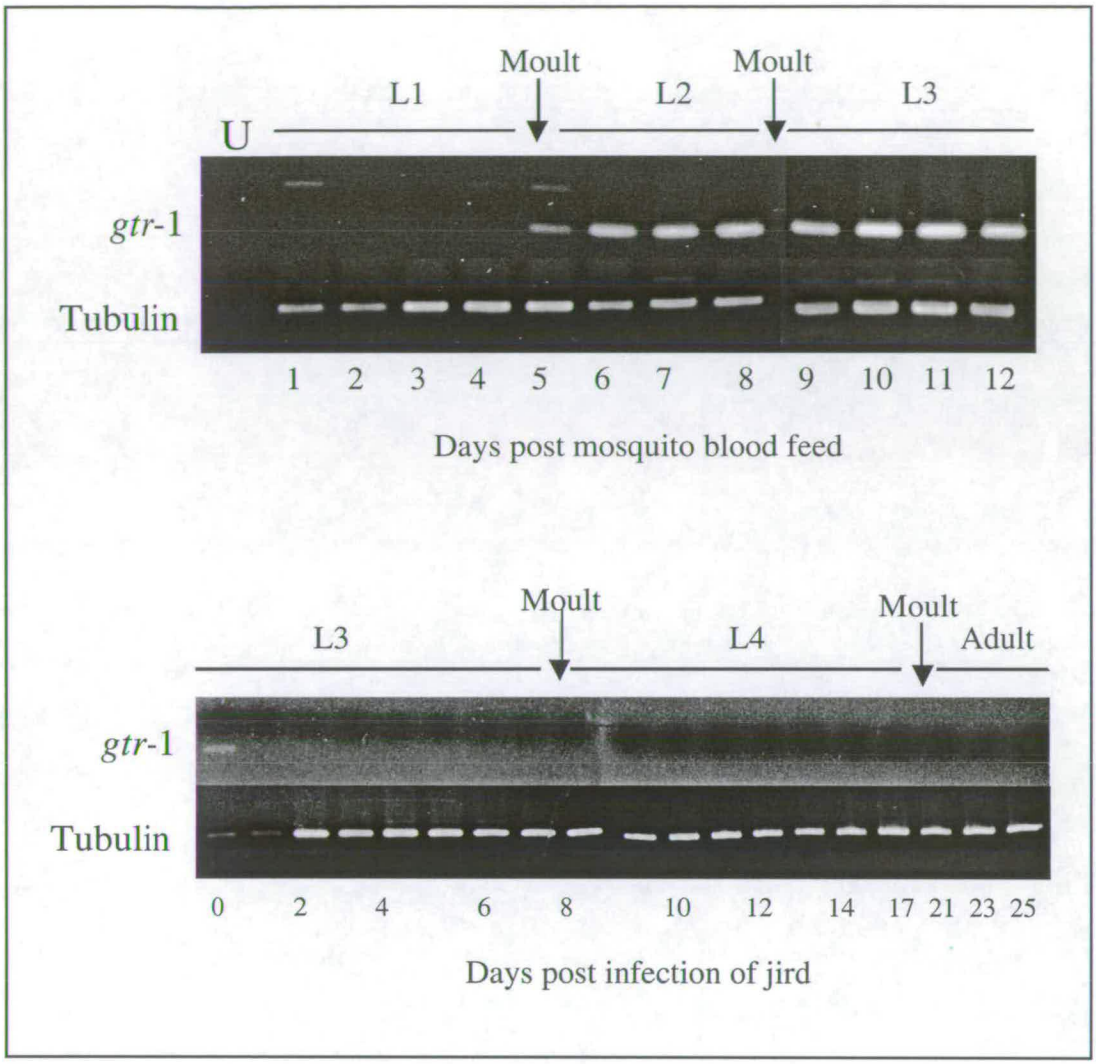


Figure 6.4: Expression analysis of the Bm-agt-1 gene. *RT-PCR was used to assess the presence of agt-1 transcripts during the complete development in the mosquito (panel A) and through the first 25 days of development in the jird (Panel B). Poly A primed 1st strand cDNA was used as a template for the PCR with agt-1-specific primers. U represents RNA extracted from mosquitoes fed on uninfected blood. Day 0 represents RNA extracted from vector-derived larvae.*

6.2.4 Expression of Bm-AGT-1 in *E. coli*

The mature *Bm-agt*-1 protein was expressed as recombinant protein in *E. coli* using the plasmid vector pET 29. The recombinant protein consisted of *agt*-1 without its signal peptide (amino acids 5 –55) along with fusions encoded by the vector: an N-terminal hexahistadine tag and a C-terminal S•Tag. The recombinant protein has a calculated molecular weight of 9.7 kDa with approximately 5.9 kDa representing the fusion partners. The expressed protein was soluble in His•Tag binding buffer but substantially more protein was extracted from the binding buffer insoluble fraction by the addition of 6M urea to the buffer. Protein was therefore purified on a metal chelate column from a 6M urea extract. Figure 6.5 shows the purification of AGT-1 from the urea-soluble extract. Once purified the protein remains soluble until dialysis into PBS is attempted. As this protein was initially required for antibody production, mice were immunised with partially precipitated protein at protein concentrations assessed on estimates taken prior to dialysis.

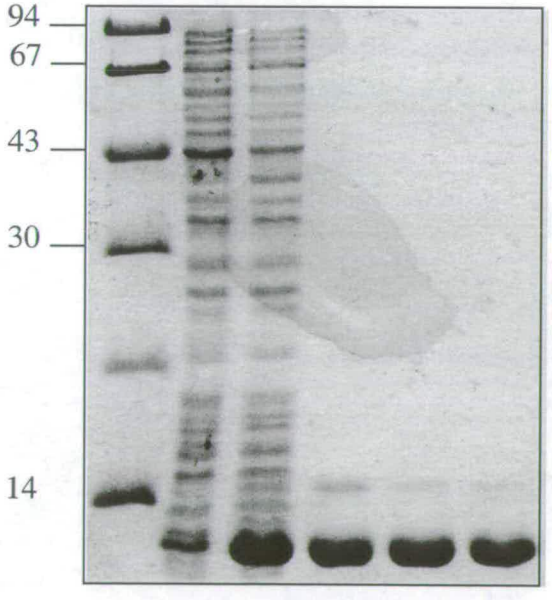


Figure 6.5: Expression and Purification of Bm-AGT-1. ***Lane 1****: low molecular weight markers.* ***Lane 2****: His•Bind binding buffer of expressing BL-21 cells.* ***Lane 3****: Urea extract of binding buffer insoluble material.* ***Lanes 4-6****: Sequential fractions eluted from the His•Bind column. Proteins were separated on a 20% SDS-PAGE gel and visualised by commassie staining.*

DISCUSSION

Interest in the *agt* sequences arose from their high glycine and tyrosine content. Tyrosine residues are able to cross-link proteins forming di- and tri-tyrosine bonds and these bonds have been reported in the insoluble residue of nematode cuticles (Fetterer and Rhoads, 1990; Fetterer et al., 1993). The role of tyrosine cross-links in the formation of insoluble of structures such as the insect eggshells (Li et al., 1996), yeast cell walls (Briza et al., 1986), sea urchin fertilisation envelopes (Kay and Shapiro, 1987) and plant call walls (Waffenschmidt et al., 1993) have also been described. The bond can be formed *in vitro* by the action of peroxidases in the presence of hydrogen peroxide. As the *agt* sequences also predict signal sequences it is possible they are also localised to the cuticle and serve as substrates for cuticular cross-linking enzymes.

Proteins rich in glycine and tyrosine residues include those forming one family of the keratin associated proteins (KAP) (Fratini et al., 1993). They are small proteins (6-9 kDa) that contribute to the insoluble scaffold of hair, nails and the outer layer of the skin. They are insoluble below pH 10 due to their unusual amino acid content. Their glycine and tyrosine residues are not found in any particular order and the family members have low nucleic acid homology arguing for the formation of the family by convergent evolution.

Chorion proteins, in contrast, are evolutionarily conserved and many contain tandemly repetitive peptides with glycine and tyrosine in their N and C terminal arms (Tsitilou et al., 1983). In this respect the nematode *agt* proteins more closely resemble chorion proteins, and indeed database searches with the *agt* protein sequences give weak hits to this family. These proteins form part of the insoluble outer shell of insect eggs and cuticle. Their gene structure and organisation also shows some similarity to the *C. elegans agt*s. Grouping of the chorion proteins by homology has allowed the designation of 3 families A, B and C. Along one chromosome the genes are found in A/B pairs and several such pairs can be found together in a 10-15 kb segment. Interestingly transcription of the A/B pairs are often

in opposite directions (Kafatos et al., 1987), similar to that found in the *C. elegans* gene cluster on cosmid B0213. Additionally chorion genes and the nematode *agt*s contain only one intron. The function of the Gly/Tyr-rich segments of the chorion proteins is not yet known although it is possible that they are involved in cross-linking as Li et al has demonstrated di- and tri- tyrosine bonds in *A. aegypti* chorion proteins (Li et al., 1996).

Proteins forming the cuticle of the adult locust, *Locusta migratoria*, are highly insoluble however the unhardened cuticle of the larval stages is more soluble. Approximately 100 proteins can be identified of which the majority are basic. N-terminal sequences of a number of these proteins showed that they contain a number of similar amino acid compositions (Højrup et al., 1986). A striking similarity is the dominant occurrence of either glycine, leucine and tyrosine or alanine and proline in certain regions.

Although not homologous a common feature of both insect cuticle and chorion proteins is the presence of regions dominated by relatively few amino acids. The nematode *agt*s however do not contain these different regions as most of their length is dominated by glycine and tyrosine, especially the *C. elegans* sequences.

These groups of proteins, although encoded for by genes of low nucleic acid homology, have a unifying property in that they tend to perform a structural or protective role. As yet there is no data on the localisation of the AGT proteins in filariae and it would be informative to know whether they localise to the insoluble zones of the cuticle. Gly/Tyr-rich proteins from plant cell walls (Waffenschmidt et al., 1993), locust cuticles (Højrup et al., 1986) and silk moth chorion (Tsitilou et al., 1983) contain the repetitive motif YGGY found in the nematode *grt* family. Addition other proteins such as *Schistosoma japonicum* egg shell proteins contain tyrosines spaced 3-4 residues apart usually with glycine used as the "spacer" amino acids (Henkle et al., 1990). This motif is common in most of the nematode *agts*. Although this consensus sequence is small, its occurrence in different insoluble, structural proteins may be an indicator of a related function. The motif may orientate

tyrosine residues within the protein ensuring their accessibility to cross-linking enzymes.

The expression pattern of *agt*-1 is unexpected, as many structural components of the cuticle described to date have been shown to be transcribed during synthesis of the new cuticle below the old cuticle. Expression of *agt*-1 is first seen once the L2 moult has been completed and is terminated after transmission, around 8 days before the L3 moult. During the eight days of expression the larvae increase in size from 0.8 mm to 1.4-1.9 mm. Although no expression of collagens is usually seen in the intermoult period (Selkirk et al., 1989) and growth is accompanied by the stretching of the preexisting cuticle (Howells and Blainey, 1983) it is possible that some components of the cuticle cannot accommodate this increase in size and are be expressed during growth. The *cut*-1 and -2 genes are the only insoluble components of the cuticle of *C. elegans* described so far. Both genes are transcribed at a time when new cuticle is being made (Lassandro et al., 1994; Sebastiano et al., 1991).

Other possible functions of *agt*s are the formation of structures other than the cuticle. Immunoelectromicroscopy with the antisera generated against the recombinant protein will help resolve these issues.

CHAPTER 7

B. malayi tumour protein homologue (*tph*-1)

INTRODUCTION

A key feature of successful parasitism is the prevention of functional antiparasitic responses in the host (Bloom, 1979; Maizels et al., 1993). During many helminth infections there is a striking block of cellular responsiveness along with a profound bias of the immune response towards a Th2-type response in both humans and animal models. The pathways underlying this phenomenon are unknown and a variety of mechanisms have been proposed to explain the domination of Th2 responses along with the down-regulation of antigen-specific Th1 responses. In a mouse model of infection with *B. pahangi* it has been shown that the L3 stimulates IL-4 transcription in the draining popliteal lymph node as early as 24 hours after infection leading to a detectable Th2-type response from day 4 post infection onwards (Osborne and Devaney, 1998). As IL-4 is both a product of and the stimulus for a Th2 response, identifying the primary source of this "early" IL-4 is a focus of recent research. Double negative ($CD4^-$, $CD8^-$) ab T cells have been identified as early IL-4 producers in *B. pahangi*-infected BALB/c mice (Osborne and Devaney, 1998). It is not clear, however, whether this population of cells responds to a parasite molecule, initiating the biasing of the response, or is responding to other sources of IL-4. Parasite molecules capable of initiating this response have not yet been identified.

Upon stimulation basophils are capable of releasing considerable amount of preformed IL-4. In a search for molecules capable of inducing histamine release from basophils, MacDonald and colleagues have shown that a molecule secreted by macrophages, termed histamine releasing factor or HRF, is capable of inducing the release of IL-4 at the same time as histamine (MacDonald, 1996; MacDonald, 1997; Schroeder et al., 1996; Schroeder et al., 1997). This process was initially described as IgE-dependent as removal of IgE from the basophil surface abrogated the effect. Moreover it relied on the presence of a particular, undefined, subset of IgE on the surface of the basophil, termed IgE^+. IgE from patients not responding to HRF were termed IgE^- although passive transfer of IgE^+ from responder patients restored the activity (Schroeder et al., 1996). Unexpectedly, HRF does not bind directly to IgE but probably signals through its own receptor on the basophil surface (Wantke et al.,

1999). Cloning of HRF identified it as the already described translationally controlled tumor protein (TCTP) (MacDonald et al., 1995) a highly conserved family of proteins described in animals, plants and yeast (Bhisutthibhan et al., 1998; Haghighat and Ruben, 1992; Pay et al., 1992; Thiele et al., 1998). The human protein is found in a wide variety of cell types (Sanchez et al., 1997). These proteins are preferentially synthesised during the early growth phase of some types of tumour, but are also expressed in normal cells. In this context the physiological function of TCTP is still not known although recent evidence suggests that it is capable of microtubule stabilisation via tubulin binding (Gachet et al., 1999).

The identification of a close homologue of TCTP amongst the abundantly expressed larval transcripts (chapter 1 and (Gregory et al., 1997)) raises the possibility that the *B. malayi* protein, termed *B m*-TPH for **t**umour **p**rotein **h**omologue, is a candidate "early IL-4" inducer.

7.2.1 Characterisation of the *Bm-tph*-1 cDNA

The *B. malayi* homologue of human histamine-releasing factor/translationally controlled tumour proteins was identified and cloned as a 750 bp abundant band amongst *trans*-spliced transcripts from the L3 stage (chapter 1 and (Gregory et al., 1997)). The cDNA predicts a protein of 181 amino acids with 32 - 67% identity to TCTP and HRFs in GenBank (Fig. 7.1). The predicted molecular weight is 20.76 kDa with an isoelectric point of 4.42. The *B. malayi* sequence contains the two signature patterns selected for TCTP by PROSITE at the Swiss Institute of Bioinformatics (available at http://www.expasy.ch/cgi-bin/nicedoc.pl?PDOC00768) and indicated in Fig. 7.2. The predicted protein lacks a conventional signal sequence (von Heijne, 1986) and does not contain any potential *N*-glycosylation motifs although other TCTPs have a potential N-glycosylation (Fig. 7.2), forming part of the TCTP-1 signature. It is not known whether this site is glycosylated within the family.

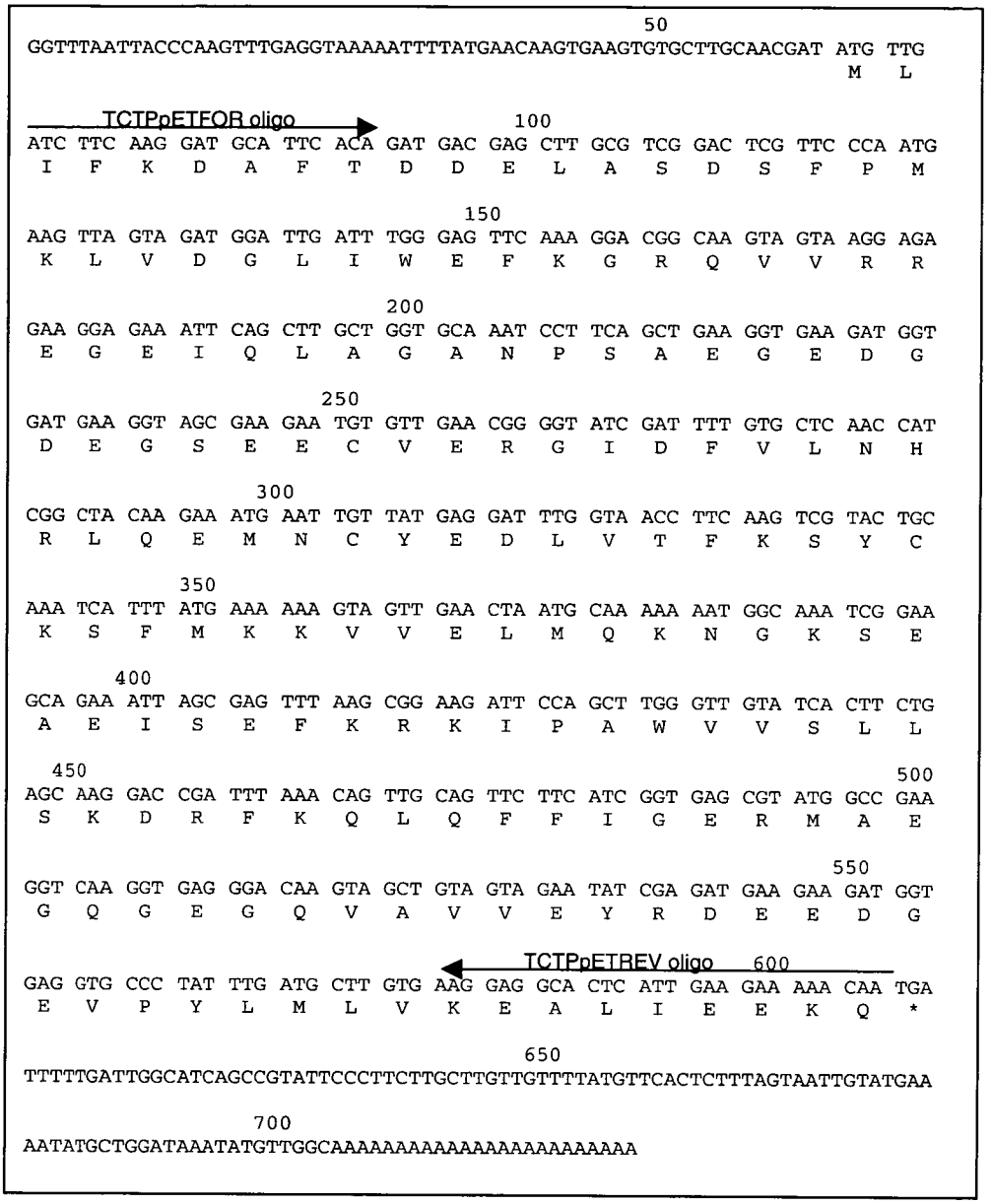


Fig. 71: Nucleotide and deduced amino acid sequence of Bm-tph-1. *The nematode spliced leader sequence SL-1 is boxed. The oligo pair indicated were used for RT-PCR and expression in pET-29.*

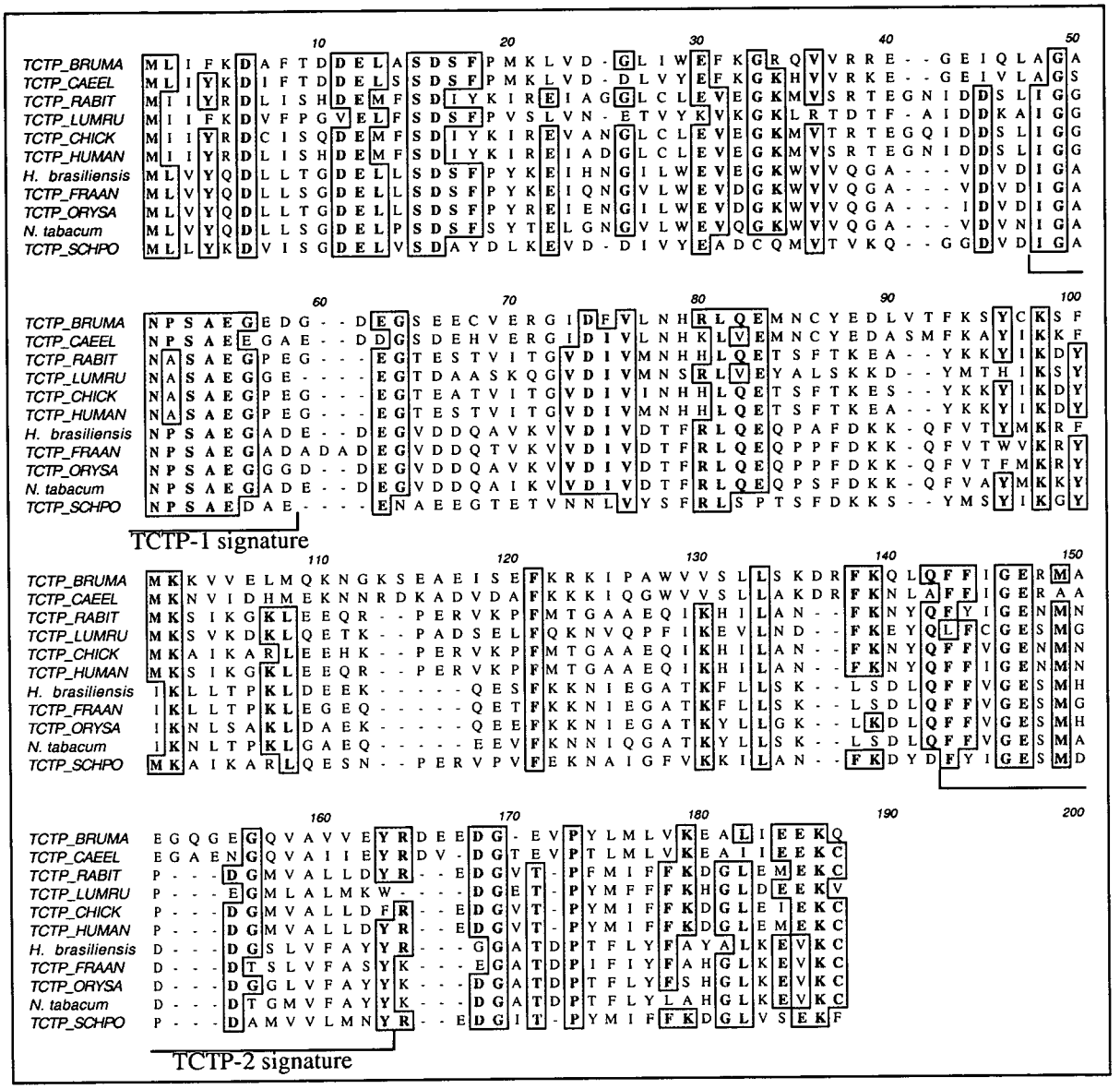


Figure 7.2.: Alignment of Bm-TPH-1 with its closest homologues. *Sequences are identified by their SwissProt identifier, if one has been assigned . SwissProt entry numbers are as follows: Bm-tph-1(TCTP_BRUMA); P90697, C. elegans (TCTP_CAEEL); Q93573, Rabbit (TCTP_RABIT); P43348, Humus earthworm(TCTP_LUMRU); O18477, Chicken (TCTP_CHICK); P43347, Human (TCTP_HUMAN); P13693, Strawberry (TCTP_FRAAN); O03992, Rice (TCTP_ORYSA); P35681 and Schizosaccharomyces pombe (TCTP_SCHPO); Q10344. GenBank accession numbers for Rubber tree (Hevea brasiliensis) and common tobacco (Nicoliana tabacum) are AF091455 and AF107842 respectively. Boundaries of the TCTP-1 and -2 signatures used to identify all TCTPs are indicated.*

7.2.2 Abundance of *Bm-tph-1* mRNA throughout the life cycle

Primers spanning the entire predicted open reading frame were used to amplify by PCR from polyA-primed first strand cDNA synthesised from parasite RNA collected during the first four weeks of infection. This time course includes the L3 - L4 moult (at around day 8) and the L4 to adult moult (at around 17 - 23 days). Although expression can be seen at most time points there are periods when expression appears to increase when compared to tubulin control (Fig. 7.2.3). In contrast to its abundant expression amongst *trans*-spliced transcripts *tph*-1 does not appear as a strong signal from vector-derived L3 cDNA. Expression was not

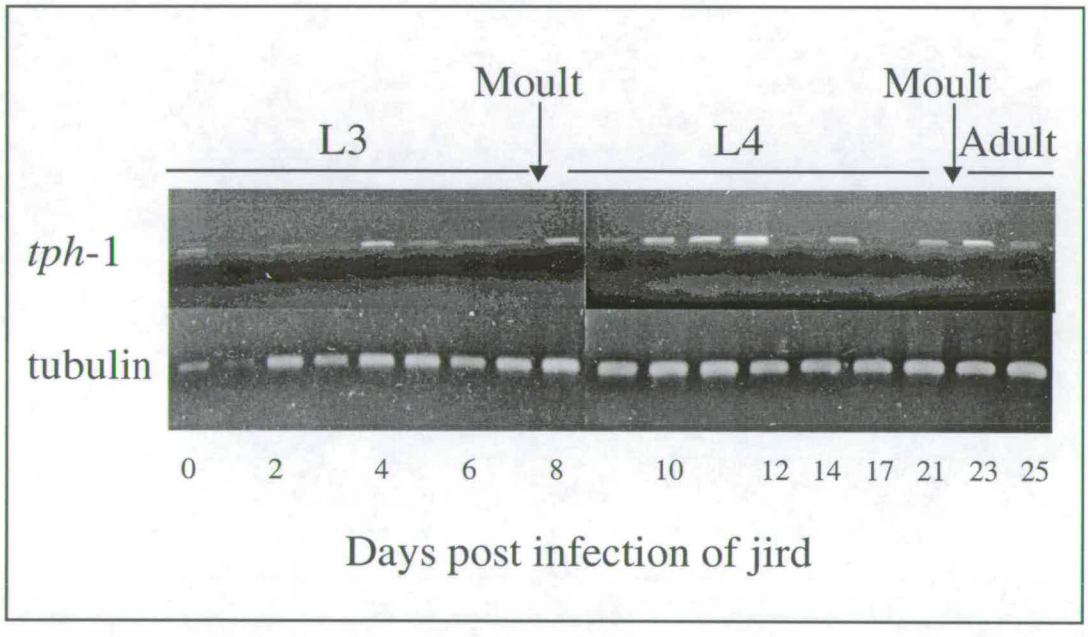


Figure 7.2.3: Expression analysis of the Bm-tph gene. *RT-PCR was used to access the presence of tph-1 transcrips during the fist 25 days of expression in the jird. PolyA-primed first strand cDNA was used as a template for PCR with tph-1 and tubulin-specific primers. Day 0 represents RNA extracted from vector-derived DNA.*

assessed during the vector stage of development. Analysis of ESTs representing *tph*-1 from the stage-specific libraries constructed by the Filarial Genome Project suggests that expression continues throughout the life cycle with peaks of expression at certain stages. Table 7.1 lists the numbers of *tph*-1 ESTs from each library, along

with the percentage of total ESTs that this represents. All libraries except the day 9 L3 have transcripts in them identified by ESTs with the adult female and mf containing the highest percentage of transcripts. This agrees well with the result of library PCRs shown in chapter 3. Libraries constructed from PCR products from SL-oligo dT amplifications are enriched compared to their conventionally constructed counterparts again suggesting that *tph*-1 is prominent amongst *trans*-spliced transcripts.

	tph-1 ESTs	Total ESTs	% of library
mf	22	3355	0.66
L2-SL*	2	615	0.33
L3	3	2761	0.11
L3-SL*	2	292	0.68
Day 6 L3	1	1545	0.06
Day 9 L3	0	259	0
L4	2	541	0.37
L4-SL*	16	1619	0.99
Young adult (day 25)	1	191	0.52
Adult Male	18	4201	0.43
Adult Female	25	3775	0.66

Table 7.1. Survey of tph-1 ESTs from B. malayi stage-specific libraries. ** denotes that libraries are constructed from SL-dT PCR products.*

7.2.3 Expression of recombinant Bm-TPH-1

The predicted full length *Bm*-TPH-1 has been expressed in *E. coli* using the plasmid pET 29. This system produces a fusion protein with an N-terminal S•Tag and a C-terminal hexahistidine tag adding approximately 5.9 kDa to the size of TPH. The calculated size of the fusion protein is 26.7 kDa although the expressed protein 1-2 kDa higher as judged by SDS-PAGE (Fig. 7.4). The fusion protein was soluble allowing the purification of the protein under non-denaturing conditions (Fig. 7.4).

Antibodies to recombinant TPH were produced in BALB/c and CBA mice. All immunised mice recognised the recombinant and a sera pool of all immunised mice was used in subsequent experiments.

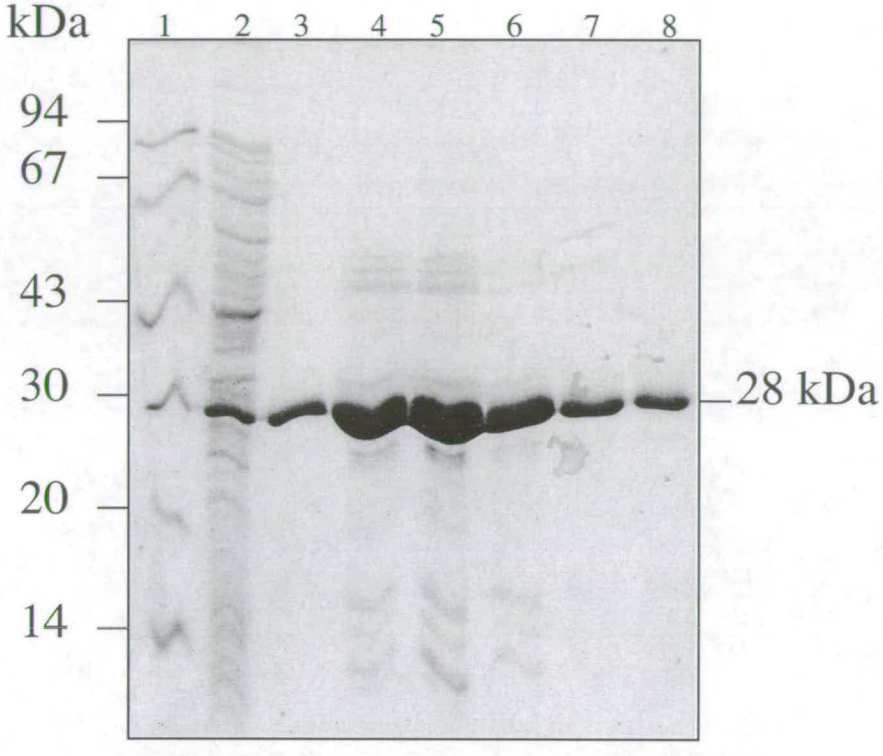


Figure 7.4: Expression and purification of Bm-TPH-1 expressed in E.coli. *Cells expressing TPH-1 from a pET-29 plasmid were solubilised in His•Bind binding buffer (**Lane 2**) and the soluble supernatant purified on a metal chelate column. **Lanes 3 - 8**: Purified fractions eluted from the column. **Lane 1** Pharmacia low molecular weight markers.*

7.2.4 Detection of *Bm*-TPH-1 in parasite extracts.

The expression of some members of the TCTP/HRF family have been shown to be under both translational and post-translational control (Bohm et al., 1989; Xu et al., 1999). It is therefore important to detect the protein as well as the mRNA at different points in the life cycle. A 25 kDa protein is detected in each stage by antisera to recombinant *Bm*-TPH-1.

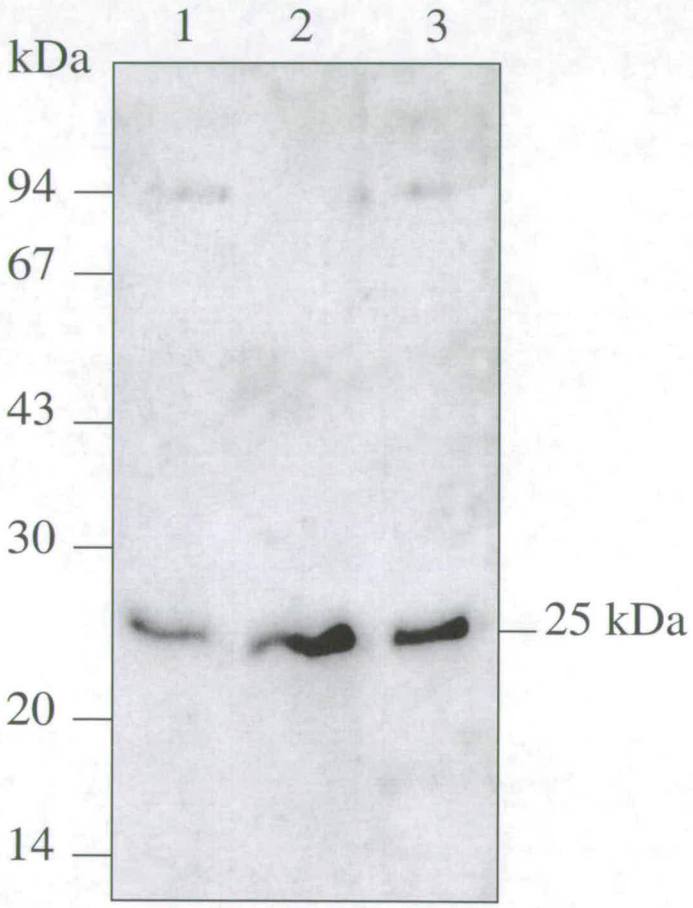


Figure 7.5: Detection of Bm-TPH in different life cycle stages. *PBS/1.5%nOG soluble extracts were prepared from mf (**Lane 1**), L3 (**Lane 2**) and adult worms (**Lane 3**). Proteins were separated on a 15% SDS-PAGE gel, transferred to nitrocellulose and probed with antisera to recombinant Bm-TPH-1.*

7.2.5 Secretion of *Bm*-TPH-1

Antisera raised against recombinant TPH-1 were used to assess whether the native protein was released from the parasite in culture. A culture of mixed sex adult worms was initially depleted of any endogenous methionine pool by incubation in methionine-free RPMI overnight. ^{35}S methionine was then added to the culture for the following 2 days at which point the culture was terminated. Without any additional processing the culture supernatant was immunoprecipitated with anti-recombinant antisera. This identified a major band of approximately 25 kDa, 5 kDa more than the predicted size.

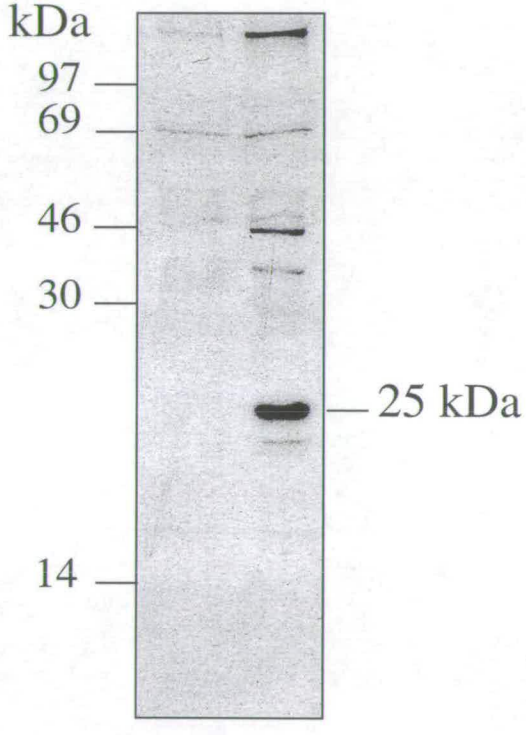


Figure 7.6. Secretion of Bm-TPH by adult parasites maintained in culture. *Adult parasites were maintained in culture for a total of 3 days. After the first 24 hours of culture 1mCi ^{35}S methionine was added. At the end of the culture period the culture supernatant was immunoprecipitated with antisera to recombinant ALT-1 (**lane 1**) and THP-1 (**Lane 2**). Antisera to recombinant ALT-1, expressed with the same fusion partners as TPH-1, was used as a negative control. Precipitated material was resolved on a 15% SDS-PAGE gel.*

DISCUSSION

A homologue of human translationally controlled tumour protein/histamine releasing factor has been identified in *B. malayi*. Originally cloned as an abundant *trans*-spliced cDNA from larval parasites (chapter 1 and (Gregory et al., 1997)). RT-PCR and analyses of ESTs from different stage-specific libraries demonstrate that the gene is expressed at most points throughout the life cycle.

The role(s) of TCTP/HRF is still unclear. It was first described as a tumour-related protein in mouse ascitic tumour and erythroleukaemic cells (Chitpatima et al., 1988; Yenofsky et al., 1983). Subsequently it has been found in a variety of tumour and normal human cells (Sanchez et al., 1997). Although expression of TCTP/HRF is not restricted to tumour cells its expression is growth-related with transcription of the gene demonstrated after mitogenic stimulation (Bohm et al., 1989; Walsh et al., 1995). Expression is controlled at both the transcriptional and translational levels with calcium involved at both of these steps (Bohm et al., 1989; Thiele et al., 1998; Xu et al., 1999). The peaks of expression in *B. malayi*, as judged by RT-PCR, do not have an absolute correlation with growth. After transmission to the jird *B. pahangi* larvae increase their length from 2 mm to 11 mm between the third moult (day 8) and the fourth moult (days 20-25) (Howells and Blainey, 1983). Expression of *Bm-tph*-1, as detected by RT-PCR, is only seen sporadically during this period (Fig. 7.3). In addition mf and mature adults, which undergo little growth, both have levels of protein detectable by western blotting. This apparent lack of a growth related expression indicates that *Bm-tph*-1 expression is unrelated to growth, or that it serves more that one function.

TCTP/HRF is also found extracellularly. It is secreted by a variety of cell types inducing the release of histamine and IL-4 from basophils (Schroeder et al., 1996). Current evidence suggests that this may involve its own receptor, as yet undefined, (Wantke et al., 1999). IL-4 and histamine are found in culture supernatants of recombinant TCTP/HRF-treated basophils within one hour. As the biasing of the immune response against filariae is thought to occur very soon after

infection (Osborne and Devaney, 1998), this timing and the presence of *Bm*-TPH-1 in larvae suggest that it is one of the candidate "Th2 inducers" in filariasis. In common with its homologues Bm-TPH-1 lacks a conventional signal sequence but is found extracellularly. Although unusual, both human and mouse TCTP/HRFs do not posses signal sequences, which may suggest an alternate route of secretion for these molecules (Muesch et al., 1990; Rubartelli et al., 1990). Secretion in the absence of a signal sequence has also been noted for OvB20, an *O. volvulus* vaccine candidate (Abdel-Wahab et al., 1996). Adult worms were chosen to illustrate secretion as this stage can be individually monitored for viability. It is however possible that TPH is released during the birthing process. The finding that adult worms actively secrete or release *Bm*-TPH-1 may indicate that a constant source of IL-4 is needed to maintain an immune response appropriate for the parasites survival.

GENERAL DISCUSSION

The rational behind these studies is that genes expressed abundantly or in a highly stage-specific manner will have important functional roles in the parasites' life cycle, and thus may provide targets for the development of novel immuno- or chemoprophylatic agents. Consequently the long-term goal of this work is the definition of function. For cDNAs that have homologues in the databases, a possible function can usually be implied. Even so there are many proteins, such as cystatins, whose precise function remains unclear. Additionally, close sequence matches in the genetically tractable nematode *C. elegans* sometimes do not prove useful in determining possible functions. For example, from this study *Bm-alt*-1 and *Bm-cpi*-1 do not appear to have close counterparts in *C. elegans* although their nearest sequence similarities are with *C. elegans* sequences. As such these genes may prove to have a role closely linked to parasitism.

This study has focused mainly on antigens expressed at the L3 stage of development as this has been demonstrated to be susceptible to immunity induced by irradiated larvae. It is hoped that by targeting incoming larvae pathology attributed to adult stages during chronic infection with lymphatic filariae will be avoided. In addition, for many of the genes identified in this study the lack of strong homology with genes from the human host provides excellent targets for chemotherapy and immunotherapy. Also notable in this study is the high number of cDNAs from vector-derived larvae encoding potential signal peptides, indicating that they may be surface and/or secreted proteins. Indeed, *Bm*-CPI-1 has been shown to be secreted by larvae in this study and ALT homologues from *Onchocerca* and *Dirofilaria* are well characterised secreted proteins. This may indicate that a large part of the parasites' activity at this point is devoted to biosynthesis of molecules that may help it establish or maintain an infection rather than be involved in general metabolic activity.

The RT-PCR technique used to identify the abundant gel is simple and economical such that it can be applied to the study of a variety nematode species. The wide diversity amongst nematodes has been illustrated in a number of recent studies that have reported sequences specific to particular species (Blaxter et al., 1996; Daub et al., 2000; Joshua and Hsieh, 1995; Maizels et al., 2000; Moore et al., 2000). Clearly the complete genome sequence of the free-living nematode will have a large impact on nematode research, however there is still a need to identify these unique sequences for use in targeted therapies and also in diagnosis of infections. An understanding of these genes would also contribute to understanding the biology underlying the diversity amongst parasitic life styles.

The abundance of *Bm-tph*-1 recommends it for further study although early analysis of its presence in the stage-specific libraries showed it to be expressed constitutively and abundantly (Williams et al., 2000). It's homology to mammalian histamine-releasing factors suggests that it may play an important role in immune modulation.

Further studies on the proteins identified in this study would benefit from recombinants produced in organisms other than *E. coli*. While both CPIs expressed as functional recombinants in *E. coli*, other molecules such as ALT, which contains eight cysteine residues potentially forming disulphide bonds, are more likely to be folded correctly when produced in other organisms, such as yeast or baculovirus. Preliminary NMR studies on the Bm-ALT-1 recombinant described in chapter 5 suggest that very little secondary structure is present. A structural model for ALT may provide a hypothesis of the function of this novel protein.

Work on the development of ALT as vaccine candidates will benefit greatly from the *L. sigmodontis* model. This filarial parasite undergoes full development in inbred laboratory mice and can be used to dissect the immune response to ALT. This study has identified two isoforms of the *alt*-2 gene differing in their intronic sequence. These may represent important polymorphic loci in filariae. It would be interesting to see if different geographical isolates also exhibit this poymorphism and whether these isoforms act as alleles.

Studies are now underway in collaboration with Benedicte Manoury and Collin Watts at the University of Dundee to investigate the effects of the *Brugia* CPIs on antigen processing in mammalian cells. As both of these proteins are secreted by parasites in culture they have the potential to interfere with host cells. *Bm*-CPI-2 has been shown to be effective in blocking cysteine protease-dependant processing in mice. The recombinant can effectively inhibit processing of tetanus toxoid by a recently discovered cysteine protease, asparaginyl endopeptidase (Manoury et al., 1998). Further work will involve testing the specificity of this inhibition and to test whether CPI-2 can block processing of the invariant chain, a process known to involve cysteine proteases. Processing of other antigens should also be tested to assess the generality of CPI-2 inhibition. Initial studies with CPI-1 show that it is not able to inhibit tetanus toxiod processing and it also fails to inhibit asparaginyl endopeptidase.

REFERENCES

Abdel-Wahab, N., Kuo, Y. M., Wu, Y., Tuan, R. S., and Bianco, A. E. (1996). OvB20, an *Onchocerca volvulus*-cloned antigen selected by differential immunoscreening with vaccination serum in a cattle model of onchocerciasis, Molecular and Biochemical Parasitology *76*, 187-199.

Abraham, D., Grieve, R. B., Mika-Grieve, M., and Siebert, B. P. (1988). Active and passive immunization of mice against larval *Dirofilaria immitis.*, Journal of Parasitology *74*, 275-282.

Abrahamson, M., Mason, R. W., Hansson, H., Buttle, D. J., Grubb, A., and Ohlsson, K. (1991). Human cystatin C: role of the N-terminal segment in the inhibition of human cysteine proteinases and in its inactivation by leucocyte elastase, Biochemical Journal *273*, 621-626.

Abrahamson, M., Ritonja, A., Brown, M. A., Grubb, A., Machleidt, W., and Barrett, A. J. (1987). Identification of the probable inhibitory reactive sites of the cysteine proteinase inhibitors human cystatin C and chicken cystatin, Journal of Biological Chemistry *262*, 9688-9694.

Adam, R., Kaltmann, B., Rudin, W., Friedrich, T., Marti, T., and Lucius, R. (1996). Identification of chitinase as the immunodominant filarial antigen recognized by sera of vaccinated rodents, Journal of Biological Chemistry *271*, 1441-1447.

Afonso, S., Romagnano, L., and Babiarz, B. (1997). The expression and function of cystatin C and cathepsin B and cathepsin L during mouse embryo implantation and placentation, Development *124*, 3415-3425.

Ahmed, S. S. (1967). Studies on the laboratory transmission of sub-periodic *Brugia malayi* and *Brugia pahangi*. I. The resistance of guinea pigs, rabbits and white mice to infection, Annals of Tropical Medicine and Parasitology *61*.

Allen, J. E., Lawrence, R. A., and Maizels, R. M. (1996). APC from mice harbouring the filarial nematode, *Brugia malayi*, prevent cellular proliferation but not cytokine production, International Immunology *8*, 143-151.

Alvarez-Fernandez, M., Barrett, A. J., Gerhartz, B., Dando, P. M., Ni, J., and Abrahamson, M. (1999). Inhibition of mammalian legumain by some cystatins is due to a novel second reactive site, Journal of Biological Chemistry *274*, 19195-19203.

Aoki, H., Akaike, T., Abe, K., Kuroda, M., Arai, S., Okamura, R., Negi, A., and Maeda, H. (1995). Antiviral effect of oryzacystatin, a proteinase inhibitor in rice, against herpes simplex virus type 1 in vitro and in vivo, Antimicrobial Agents and Chemotherapy *39*, 846-849.

Apfel, H., Eisenbeiss, W. F., and Meyer, T. F. (1992). Changes in the surface composition after transmission of *Acanthocheilonema viteae* third stage larvae into the jird, Molecular and Biochemical Parasitology *52*, 63-74.

Auerswald, E. A., Genenger, G., Assfalg-Machleidt, I., Machleidt, W., Engh, R. A., and Fritz, H. (1992). Recombinant chicken egg white cystatin variants of the QLVSG region, European Journal of Biochemistry *209*, 837-845.

Bailly, S., Israel, N., Fay, M., Gougerot-Pocidalo, M. A., and Duff, G. W. (1996). An intron polymorphic repeat sequence modulates interleukin-1 alpha gene regulation, Molecular Immunology *33*, 999-1006.

Barna, J. B., and Kew, R. R. (1995). Inhibition of neutrophil chemotaxis by protease inhibitors. Differential effect of inhibitors of serine and thiol proteases, Inflammation *19*, 561-74.

Barrett, A. J. (1987). The cystatins: a new class of peptidase inhibitors, Trends in Biochemical Sciences *12*, 193-196.

Barrett, A. J., Fritz, H., Grubb, A., Isemura, S., Jarvinen, M., Katunuma, N., Machleidt, W., Muller-Esterl, W., Sasaki, M., and Turk, V. (1986). Nomenclature and classification of the proteins homologous with the cysteine-proteinase inhibitor chicken cystatin, Biochemical Journal *236*, 312.

Bhisutthibhan, J., Pan, X. Q., Hossler, P. A., Walker, D. J., Yowell, C. A., Carlton, J., Dame, J. B., and Meshnick, S. R. (1998). The *Plasmodium falciparum* translationally controlled tumor protein homolog and its reaction with the antimalarial drug artemisinin, Journal of Biological Chemistry *273*, 16192-16198.

Bianco, A. E., Robertson, B. D., Kuo, Y.-M., Townson, S., and Ham, P. J. (1990). Developmentally regulated expression and secretion of a polymorphic antigen by *Onchocerca* infective stage larvae, Molecular and Biochemical Parasitology *39*, 203-212.

Bianco, A. E., Wu, Y., and Jenkins, R. E. (1995). *Onchocerca* spp: A "family" of secreted acidic proteins expressed in infective larave in blackflies, Experimental Parasitology *81*, 344-354.

Bird, A. F., and Bird, J. (1991). The Stucture of Nematodes, Academic Press).

Blaxter, M. L. (1995). The Filarial Genome Project, Parasitology Today *11*, 811-812.

Blaxter, M. L., and Liu, L. X. (1996). Nematode spliced leaders - function, evolution and utility, International Journal of Parasitology *26*, 1025-1033.

Blaxter, M. L., Raghavan, N., Ghosh, I., Guiliano, D., Lu, W., Williams, S. A., Slatko, B., and Scott, A. L. (1996). Genes expressed in *Brugia malayi* infective larvae, Molecular and Biochemical Parasitology *77*, 77-93.

Blaxter, M. L., and Robertson, W. M. (1998). The cuticle. In The Physiology and Biochemistry of Free-living and Plant -parasitic nematodes, R. N. Perry, and D. J. Wright, eds. (CAB International).

Bloom, B. R. (1979). Games parasites play: how parasites evade immune survailance, Nature *279*, 21-26.

Blumenthal, T., and Steward, K. (1997). RNA processing and gene structure. In *C. elegans* II, D. L. Riddle, T. Blumenthal, B. J. Meyer, and J. R. Priess, eds. (Cold Spring Harbor Laboratory Press), pp. 117-145.

Bode, W., Brzin, J., and Turk, V. (1985). Crystallization of chicken egg white cystatin, a low molecular weight protein inhibitor of cysteine proteinases, and preliminary X-ray diffraction data, Journal of Molecular Biology *181*, 331-332.

Bode, W., Engh, R., Musil, D., Thiele, U., Huber, R., Karshikov, A., Brzin, J., Kos, J., and Turk, V. (1988). The 2.0 Å X-ray crystal structure of chicken egg white cystatin and its possible mode of interaction with cysteine proteinases, EMBO Journal *7*, 2593-2599.

Bohm, H., Benndorf, R., Gaestel, M., Gross, B., Nurnberg, P., Kraft, R., Otto, A., and Bielka, H. (1989). The growth-related protein P23 of the Ehrlich ascites tumor: translational control, cloning and primary structure, Biochemistry International *19*, 277-286.

Briza, P., Breitenbach, M., Ellinger, A., and Segall, J. (1990). Isolation of two developmentally regulated genes involved in spore wall maturation in *Saccharomyces cerevisiae*, Genes and Development *10*, 1775-1789.

Briza, P., Winkler, G., Kalchhauser, H., and Breitenbach, M. (1986). Dityrosine is a prominent component of the yeast ascospore wall. A proof of its structure, Journal of Biological Chemistry *261*, 4288-4294.

Brown, W. M., and Dziegielewska, K. M. (1997). Friends and relations of the cystatin superfamily-new members and their evolution, Protein Science *6*, 5-12.

Bundy, D. A. P., Grenfell, B. T., and Rajagopalan, P. K. (1991). Immunoepidemiology of lymphatic filariasis: the relationship between infection and disease, Immunology Today *12*, A71-A75.

Carlow, C. K. S., and Phillip, M. (1987). Protective immunity to *Brugia malayi* larvae in BALB/c mice: potential of this model for the indentification of protective antigens, American Journal of Tropical Medicine and Hygiene *37*, 597-604.

Chapman, H. A. (1998). Endosomal proteolysis and MHC class II function, Current Opinion in Immunology *10*, 93-102.

Chen, J. M., Dando, P. M., Rawlings, N. D., Brown, M. A., Young, N. E., Stevens, R. A., Hewitt, E., Watts, C., and Barrett, A. J. (1997). Cloning, isolation, and characterization of mammalian legumain, an asparaginyl endopeptidase, Journal of Biological Chemistry *272*, 8090-8098.

Chikilian, M. L., Bradley, T. J., Nayar, J. K., and Knight, J. W. (1994). Ultrastructural comparison of extracellular and intracellular encapsulation of *Brugia malayi* in *Anopheles quadrimaculatus*, Journal of Parasitology *80*, 133-140.

Chitpatima, S. T., Makrides, S., Bandyopadhyay, R., and Brawerman, G. (1988). Nucleotide sequence of a major messenger RNA for a 21 kilodalton polypeptide that is under translational control in mouse tumor cells, Nucleic Acids Research *16*, 2350.

Chusattayanond, W., and Denham, D. A. (1986). Attempted vaccination of jirds (*Meriones unguiculatus*) against *Brugia pahangi* with radiation attenuated infective larvae, Journal of Helminthology *60*, 149-155.

Chusattayond, W., and Denham, D. A. (1984). Induction of host resistance to *Brugia pahangi* in jirds (*Meriones unguiculatus*) protected by chemoprophylaxis., Journal of Helminthology *58*, 245-249.

Cimerman, N., Kosorok, M. D., Korant, B. D., Turk, B., and Turk, V. (1996). Characterization of cystatin C from bovine parotid glands: cysteine proteinase inhibition and antiviral properties, Biological Chemistry Hoppe Seyler *377*, 19-23.

Collins, A. R., and Grubb, A. (1998). Cystatin D, a natural salivary cysteine protease inhibitor, inhibits coronavirus replication at its physiologic concentration, Oral Microbiology Immunology *13*, 59-61.

Cookson, E., Blaxter, M. L., and Selkirk, M. E. (1992). Identification of the major soluble cuticular glycoprotein of lymphatic filarial nematode parasites (gp29) as a secretory homolog of glutathione peroxidase, Proceedings of the National Academy of Sciences USA *89*, 5837-5841.

Coulibaly, S., Schwihla, H., Abrahamson, M., Albini, A., Cerni, C., Clark, J. L., Ng, K. M., Katunuma, N., Schlappack, O., Glossl, J., and Mach, L. (1999). Modulation of invasive properties of murine squamous carcinoma cells by heterologous expression of cathepsin B and cystatin C, International Journal of Cancer *83*, 526-531.

Cox, J. L., Sexton, P. S., Green, T. J., and Darmani, N. A. (1999). Inhibition of B16 melanoma metastasis by overexpression of the cysteine proteinase inhibitor cystatin C, Melanoma Research *9*, 369-374.

Dalton, J. P., Hola-Jamriska, L., and Brindley, P. J. (1995). Asparaginyl endopeptidase activity in adult *Schistosoma mansoni*, Parasitology *111*, 575-80.

Dando, P. M., Fortunato, M., Smith, L., Knight, C. G., McKendrick, J. E., and Barrett, A. J. (1999). Pig kidney legumain: an asparaginyl endopeptidase with restricted specificity, Biochemical Journal *339*, 743-749.

Daub, J., Loukas, A., Pritchard, D. I., and Blaxter, M. (2000). A survey of genes expressed in adults of the human hookworm, *Necator americanus*, Parasitology *120*, 171-184.

Day, K. P. (1991). The endemic normal in filariasis: a static concept., Parasitology Today *7*, 341-343.

Day, K. P., Gregory, W. F., and Maizels, R. M. (1991a). Age-specific acquisition of immunity to infective larvae in a Bancroftian filariasis endemic area of Papua New Guinea., Parasite Immunology *13*, 277-290.

Day, K. P., Grenfell, B., Spark, R., Kazura, J. W., and Alpers, M. P. (1991b). Age specific patterns of change in the dynamics of *Wuchereria bancrofti* infection in Papua New Guinea, American Journal of Tropical Medicine and Hygiene *44*, 518-527.

Denham, D. A., and Fletcher, C. (1987). The cat infected with *Brugia pahangi* as a model of human filariasis. In Fiariasis, D. Evered, and S. Clark, eds. (Chichester, John Wiley and Sons), pp. 225-235.

Denham, D. A., and McGreevy, P. B. (1977). Brugian filariasis: Epidemiological and Experimental Studies. In Advances in Parasitology, B. Dawes, ed. (Academic Press), pp. 243-307.

Denham, D. A., Medeiros, F., Baldwin, C., Kumar, H., Midwinter, I. C., Birch, D. W., and Smail, A. (1992). Repeated infection of cats with *Brugia pahangi*: parasitological observations, Parasitology *104*, 415-420.

Deussing, J., Roth, W., Saftig, P., Peters, C., Ploegh, H. L., and Villadangos, J. A. (1998). Cathepsins B and D are dispensable for major histocompatibility complex class II-mediated antigen presentation, Proceedings of the National Academy of Sciences USA *95*, 4516-4521.

Dissanayake, S., Xu, M., Nkenfou, C., and Piessens, W. F. (1993). Molecular cloning and serological characterization of a *Brugia malayi* pepsin inhibitor homolog, Molecular and Biochemical Parasitology *62*, 143-146.

Dreyer, G., Addiss, D., Noroes, J., Amaral, F., Rocha, A., and Coutinho, A. (1996). Ultrasonographic assessment of the adulticidal efficacy of repeat high- dose ivermectin in bancroftian filariasis, Tropical Medicine and International Health *1*, 427-432.

Dreyer, G., Addiss, D., Santos, A., Figueredo-Silva, J., and Noroes, J. (1998). Direct assessment in vivo of the efficacy of combined single-dose ivermectin and diethylcarbamazine against adult *Wuchereria bancrofti*, Transactions of the Royal Society of Tropical Medicine and Hygiene *92*, 219-222.

Eisenbeiss, W. F., Apfel, H., and Meyer, T. F. (1994). Protective immunity linked with a distinct developmental stage of a filarial parasite, Journal of Immunology *152*, 735-742.

Estrada, S., Nycander, M., Hill, N. J., Craven, C. J., Waltho, J. P., and Bjork, I. (1998). The role of Gly-4 of human cystatin A (stefin A) in the binding of target proteinases. Characterization by kinetic and equilibrium methods of the interactions of cystatin A Gly-4 mutants with papain, cathepsin B, and cathepsin L, Biochemistry *37*, 7551-7560.

Evans, D., Zorio, D., MacMorris, M., Winter, C. E., Lea, K., and Blumenthal, T. (1997). Operons and SL2 trans-splicing exist in nematodes outside the genus *Caenorhabditis*, Proceedings of the National Academy of Sciences USA *94*, 9751-9756.

Ewert, A., and El Bihari, S. (1971). Rapid recovery of *Brugia malayi* larvae following experimental infection of cats., Transactions of the Royal Society of Tropical Medicine and Hygiene *65*, 364-368.

Ewert, A., and Ho, B. C. (1967). The fate of *Brugia pahangi* larvae immediately after feeding by infective vectormosquitoes., Transactions of the Royal Society of Tropical Medicine and Hygiene *61*.

Fetterer, R. H., and Rhoads, M. L. (1990). Tyrosine-derived cross-linking amino acids in the sheath of *Heamonchus contortus* infective larvae, Journal of Parasitology *76*, 619-624.

Fetterer, R. H., Rhoads, M. L., and Urban, J. F. (1993). Synthesis of tyrosine-derived cross-links in *Ascaris suum* cuticular proteins, Journal of Parasitology *79*, 160-166.

Frank, G. R., and Grieve, R. B. (1991). Metabolic labeling of *Dirofilaria immitis* third- and fourth-stage larvae and their excretory-secretory products., Journal of Parasitology *77*, 950-956.

Frank, G. R., and Grieve, R. B. (1996). Purification and characterization of three larval excretory-secretory proteins from *Dirofilaria immitis*, Molecular and Biochemical Parasitology *75*, 221-229.

Frank, G. R., Tripp, C. A., and Grieve, R. B. (1996). Molecular cloning of a developmentally regulated protein isolated from excretory-secretory products of larval *Dirofilaria immitis*, Molecular and Biochemical Parasitology *75*, 231-240.

Fratini, A., Powell, B. C., and Rogers, G. E. (1993). Sequence, expression, and evolutionary conservation of a gene encoding a glycine/tyrosine-rich keratin-associated protein of hair, Journal of Biological Chemistry *268*, 4511-4518.

Freedman, D. O., Nutman, T. B., and Ottesen, E. A. (1989). Protective immunity in Bancroftian filariasis. Selective recognition of a 43-kD larval stage antigen by infection-free individuals in an endemic area., Journal of Clinical Investigation *83*, 14-22.

Fuhrman, J. A., Lane, W. S., Smith, R. F., Piessens, W. F., and Perler, F. B. (1992). Transmission-blocking antibodies recognize microfilarial chitinase in brugian lymphatic filariasis, Proceedings of the National Academy of Sciences USA *89*, 1548-1552.

Gachet, Y., Tournier, S., Lee, M., Lazaris-Karatzas, A., Poulton, T., and Bommer, U. A. (1999). The growth-related, translationally controlled protein P23 has properties of a tubulin binding protein and associates transiently with microtubules during the cell cycle, Journal of Cell Science *112*, 1257-1271.

Garcia, J. A., Cervera, M. T., Riechmann, J. L., and Lopez-Otin, C. (1993). Inhibitory effects of human cystatin C on plum pox potyvirus proteases, Plant Molecular Biology *22*, 697-701.

Gelband, H. (1994). Diethylcarbamazine salt in the control of lymphatic filariasis, American Journal of Tropical Medicine and Hygiene *50*, 655-662.

Gems, D. H., Ferguson, C. J., Robertson, B. D., Nieves, R., Page, A. P., Blaxter, M. L., and Maizels, R. M. (1994). An abundant, *trans*-spliced mRNA from *Toxocara canis* infective larvae encodes a 26-kDa protein with homology to phosphatidylethanolamine-binding proteins., Journal of Biological Chemistry *270*, 18517-18522.

Gems, D. H., and Maizels, R. M. (1996). An abundantly expressed mucin-like protein from *Toxocara canis* infective larvae: The precursor of the larvae surface coat glycoproteins, Proceedings of the National Academy of Sciences USA *93*, 1665-1670.

Genenger, G., Lenzen, S., Mentele, R., Assfalg-Machleidt, I., and Auerswald, E. A. (1991). Recombinant Q53E- and Q53N--chicken egg white cystatin variants inhibit papain, actinidin and cathepsin B, Biomedica Biochimica Acta *50*, 621-625.

Ghiso, J., Jensson, O., and Frangione, B. (1986). Amyloid fibrils in hereditary cerebral hemorrhage with amyloidosis of Icelandic type is a variant of gamma-trace basic protein (cystatin C), Proceedings of the National Academy of Sciences USA *83*, 2974-2978.

Ghosh, I., Eisinger, S. W., Raghavan, N., and Scott, A. L. (1998). Thioredoxin peroxidases from *Brugia malayi*, Molecular and Biochemical Parasitology *91*, 207-220.

Greene, B. M. (1987). Primate model for onchocerciasis research. In Filariasis, pp. 236-43.

Gregory, W. F., Blaxter, M. L., and Maizels, R. M. (1997). Differentially expressed, abundant *trans*-spliced cDNAs from larval *Brugia malayi*, Molecular and Biochemical Parasitology *87*, 85-95.

Grencis, R. K. (1996). T cell and cytokine basis of host variability in response to intestinal namotode infections, Parasitology *112*, S31-S37.

Grieve, R. B., Abraham, D., Mika-Grieve, M., and Siebert, B. P. (1988). Induction of protective immunity in dogs to infection with *Dirofilaria immitis* using chemically abbreviated infections., American Journal of Tropical Medicine and Hygiene *39*, 373-379.

Guenette, S., Prichard, R. K., and Matlashewski, G. (1992). Identification of a novel *Brugia pahangi* beta-tubulin gene (beta 2) and a 22-nucleotide spliced leader sequence on beta 1-tubulin mRNA, Molecular and Biochemical Parasitology *50*, 275-284.

Haghighat, N. G., and Ruben, L. (1992). Purification of novel calcium binding proteins from *Trypanosoma brucei*: properties of 22-, 24- and 38-kilodalton proteins, Molecular and Biochemical Parasitology *51*, 99-110.

Halfon, S., Ford, J., Foster, J., Dowling, L., Lucian, L., Sterling, M., Xu, Y., Weiss, M., Ikeda, M., Liggett, D., *et al.* (1998). Leukocystatin, A new class II cystatin expressed selectively by hematopoietic cells, Journal of Biological Chemistry *273*, 16400-16408.

Hall, A., Dalboge, H., Grubb, A., and Abrahamson, M. (1993). Importance of the evolutionarily conserved glycine residue in the N- terminal region of human cystatin C (Gly-11) for cysteine endopeptidase inhibition, Biochemical Journal *291*, 123-129.

Hall, A., Hakansson, K., Mason, R. W., Grubb, A., and Abrahamson, M. (1995). Structural basis for the biological specificity of cystatin C: Identification of leucine 9 in the N-terminal binding region as a selectivity-conferring residue in the inhibition of mammalian cysteine peptidases, Journal of Biological Chemistry *270*, 5115-5121.

Hartmann, S., Adam, R., Marti, T., Kirsten, C., Seidinger, S., and Lucius, R. (1997a). A 41-kDa antigen of the rodent filaria *Acanthocheilonema viteae* with homologies to tropomyosin induces host-protective immune responses, Parasitology Research *83*, 390-393.

Hartmann, S., Kyewski, B., Sonnenburg, B., and Lucius, R. (1997b). A filarial cysteine protease inhibitor down-regulates T cell proliferation and enhances interleukin-10 production, European Journal of Immunology *27*, 2253-2260.

Haugen, A., Mann, D., Murray, C., Weston, A., and Willey, J. C. (1989). Interleukin-1 alpha gene intron containing variable repeat region coding for the SP1 transcriptional factor recognition sequence is polymorphic, Molecular Carcinogenesis *2*, 68-71.

Hayashi, Y., Noda, K., Shirasaka, A., Nogami, S., and Nakamura, M. (1984). Vaccination of BALB/c mice against *Brugia malayi* and *B. pahangi* with larvae attenuated by gamma irradiation, Japan Journal Experimental Medicine *54*, 177-181.

Henkle, K. J., Cook, G. A., Foster, L. A., Engman, D. M., Bobek, L. A., Cain, G. D., and Donelson, J. E. (1990). The gene family encoding eggshell proteins of *Schistosoma japonicum*, Molecular and Biochemical Parasitology *42*, 69-82.

Hoerauf, A., Nissen-Pahle, K., Schmetz, C., Henkle-Duhrsen, K., Blaxter, M. L., Buttner, D. W., Gallin, M. Y., Al-Qaoud, K. M., Lucius, R., and Fleischer, B. (1999). Tetracycline therapy targets intracellular bacteria in the filarial nematode *Litomosoides sigmodontis* and results in filarial infertility, Journal of Clinical Investigation *103*, 11-18.

Højrup, P., Andersen, S. O., and Roepstorff, P. (1986). Isolation, characterization, and N-terminal sequence studies of cuticular proteins from the migratory locust, *Locusta migratoria*, European Journal of Biochemistry *154*, 153-159.

Hong, X., Bouvier, J., Wong, M. M., Yamagata, G. Y., and McKerrow, J. H. (1993). *Brugia pahangi*: identification and characterization of an aminopeptidase associated with larval molting, Experimental Parasitology *76*, 127-133.

Howells, R. E., and Blainey, L. J. (1983). The moulting process and the phenomenon of intermoult growth in the filarial nematode *Brugia pahangi*, Parasitology *87*, 493-505.

Hussain, R., Grögl, M., and Ottesen, E. A. (1987). IgG antibody subclasses in human filariasis. Differential subclass recognition of parasite antigens correlates with different clinical manifestations of infection., Journal of Immunology *139*, 2794-2798.

Ibrahim, M. S., Richie, T. L., and Scott, A. L. (1992). Surface-associated antigens of *Brugia malayi* L2 and L3 parasites during vector-stage development, Molecular and Biochemical Parasitology *52*, 97-110.

Ismail, M. M., Jayakody, R. L., Weil, G. J., Nirmalan, N., Jayasinghe, K. S., Abeyewickrema, W., Rezvi Sheriff, M. H., Rajaratnam, H. N., Amarasekera, N., de Silva, D. C., *et al.* (1998). Efficacy of single dose combinations of albendazole, ivermectin and diethylcarbamazine for the treatment of bancroftian filariasis, Transactions of the Royal Society of Tropical Medicine and Hygiene *92*, 94-97.

Jenkins, R. E., Taylor, M. J., Gilvary, N., and Bianco, A. E. (1996). Characterization of a secreted antigen of Onchocerca volvulus with host- protective potential, Parasite Immunology *18*, 29-42.

Joeseph, G. T., Huima, T., and Lustigman, S. (1998). Characterization of an *Onchocerca volvulus* L3-specific antigen, *Ov*-ALT-1., Molecular and Biochemical Parasitology *96*, 177-183.

Johnstone, I. L., and Barry, J. D. (1996). Temporal reiteration of a precise gene expression pattern during nematode development, EMBO Journal *15*, 3633-3639.

Joshua, G. W., and Hsieh, C. Y. (1995). Stage-specifically expressed genes of *Angiostrongylus cantonensis*: identification by differential display, Molecular and Biochemical Parasitology *71*, 285-289.

Jovanovic, M., Sokolic, N., Movsesijan, M., and Cuperlovic, K. (1965). Immunisation of sheep with the irradiated larvae of *Dictocaulus filaria*, British Veterinary Journal *121*, 119-131.

Kafatos, F. C., Spoerel, N., Mitsialis, S. A., Nguyen, H. T., Romano, C., Lingappa, J. R., Mariani, B. D., Rodakis, G. C., Lecanidou, R., and Tsitilou, S. G. (1987). Developmental control and evolution in the chorion gene families of insects, Advances in Genetics *24*, 223-242.

Katunuma, N., and Kominami, E. (1995). Structure, properties, mechanisms, and analysis of cysteine protease inhibitors: cystatins and E64 derivitives. In Methods in Enzymology, pp. 382-397.

Kaushal, N. A., Hussain, R., Nash, T. R., and Ottesen, E. A. (1982). Identification and characterisation of excretory-secretory products of *Brugia malayi*, adult filarial parasites, Journal of Immunology *129*, 338-343.

Kay, E. S., and Shapiro, B. M. (1987). Ovoperoxidase assembly into the sea urchin fertilisation envelope and dityrosine crosslinking, Developmental Biology *121*, 325-334.

Kazura, J., Greenberg, J., Perry, R., Weil, G., Day, K., and Alpers, M. (1993). Comparison of single-dose diethylcarbamazine and ivermectin for treatment of bancroftian filariasis in Papua New Guinea, American Journal of Tropical Medicine and Hygiene *49*, 804-811.

Kazura, J. W., Cicirello, H., and McCall, J. W. (1986). Induction of protection against *Brugia malayi* infection in jirds by microfilarial antigens, Journal of Immunology *136*, 1422-1426.

Kazura, J. W., Maroney, P. A., Pearlman, E., and Nilsen, T. W. (1990). Protective efficacy of a cloned *Brugia malayi* antigen in a mouse model of microfilaremia, Journal of Immunology *145*, 2260-2264.

Kennedy, M. W., Allen, J. E., Wright, A. S., McCruden, A. B., and Cooper, A. (1995). The gp15/400 polyprotein antigen of Brugia malayi binds fatty acids and retinoids, Molecular and Biochemical Parasitology *71*, 41-50.

Kläger, S. L., Hagen, H.-E., and Bradley, J. E. (1999). Effects of an *Onchocerca*-derived cysteine protease inhibitor on microfilariae in thier simuliid vector, Parasitology *118*, 305-310.

Korant, B. D., Brzin, J., and Turk, V. (1985). Cystatin, a protein inhibitor of cysteine proteases alters viral protein cleavages in infected human cells, Biochemical Biophysical Research Communications *127*, 1072-1076.

Kos, J., Stabuc, B., Cimerman, N., and Brunner, N. (1998). Serum cystatin C, a new marker of glomerular filtration rate, is increased during malignant progression, Clinical Chemistry *44*, 2556-2557.

Kuopio, T., Kankaanranta, A., Jalava, P., Kronqvist, P., Kotkansalo, T., Weber, E., and Collan, Y. (1998). Cysteine proteinase inhibitor cystatin A in breast cancer, Cancer Research *58*, 432-436.

Kurniawan, A., Yazdanbakhsh, M., van Ree, R., Aalberse, R., Selkirk, M. E., Partono, F., and Maizels, R. M. (1993). Differential expression of IgE and IgG4 specific antibody responses in asymptomatic and chronic human filariasis, Journal of Immunology *150*, 3941-3950.

Kurniawan-Atmadja, A., Sartono, E., Partono, F., Yazdanbakhsh, M., and Maizels, R. M. (1998). Antibody responses to filarial infective larvae are not dominated by the IgG4 isotype, Parasite Immunology *20*, 9-17.

Kwan-Lim, G.-E., Gregory, W. F., Selkirk, M. E., Partono, F., and Maizels, R. M. (1998). Secreated antigens of filarial nematodes: a survay and characterisation of *in vitro* excreated/secreated products of adult *Brugia malayi*., Parasite Immunology *11*, 629-654.

Laing, A. B. G., Edison, J. F. B., and Wharton, R. H. (1961). Studies on filariasis in Malayi: further experiments on the transmission of *Brugia malayi* and *Wuchereria bancrofti*, Annals of Tropical Medicine and Parasitology *5*, 86-92.

Lal, R. B., and Ottesen, E. A. (1988). Characterization of stage-specific antigens of infective larvae of the filarial parasite *Brugia malayi*, Journal of Immunology *140*, 2032-2038.

Lassandro, F., Sebastiano, M., Zei, F., and Bazzicalupo, P. (1994). The role of dityrosine formation in the crosslinking of CUT-2, the product of a second cuticlin gene of Caenorhabditis elegans, Molecular and Biochemical Parasitology *65*, 147-59.

Lawrence, R. A. (1996). Lymphatic filariasis: what mice can tell us., Parastology Today *12*, 267-271.

Le Goff, L., Marechal, P., Petit, G., Taylor, D. W., Hoffman, W., and Bain, O. (1997). Early reduction of the challenge recovery rate following immunization with irradiated infective larvae in a filarial mouse system., Tropical Medicine and International Health *2*, 1170-1174.

Lerner, U. H., Johansson, L., Ranjso, M., Rosenquist, J. B., Reinholt, F. P., and Grubb, A. (1997). Cystatin C, and inhibitor of bone resorption produced by osteoblasts, Acta Physiologica Scandinavia *161*, 81-92.

Li, B.-W., Chandrashekar, R., Alvarez, R. M., Liftis, F., and Weil, G. J. (1991). Identification of paramyosin as a potential protective antigen against *Brugia malayi* infection in jirds., Molecular and Biochemical Parasitology *49*, 315-324.

Li, B. W., Chandrashekar, R., and Weil, G. J. (1993). Vaccination with recombinant filarial paramyosin induces partial immunity to *Brugia malayi* infection in jirds, Journal of Immunology *150*, 1881-1885.

Li, B. W., Zhang, S., Curtis, K. C., and Weil, G. J. (1999). Immune responses to *Brugia malayi* paramyosin in rodents after DNA vaccination, Vaccine *18*, 76-81.

Li, J., Hodgeman, B. A., and Christensen, B. M. (1996). Involvement of peroxidase in chorion hardening in *Aedes aegypti*, Insect Biochemistry and Molecular Biology *26*, 309-317.

Lindahl, P., Ripoll, D., Abrahamson, M., Mort, J. S., and Storer, A. C. (1994). Evidence for the interaction of valine-10 in cystatin C with the S2 subsite of cathepsin B, Biochemistry *33*, 4384-4392.

Lucius, R., Ruppel, A., and Diesfeld, H. J. (1986). *Dipetalonema viteae* : Resistance in *Meriones unguiculatus* with multiple infections of stage-3 larvae, Experimental Parasitology *62*, 237-246.

Lucius, R., Textor, G., Kern, A., and Kirstern, C. (1991). *Acanthocheilonema viteae* : Vaccination of jirds with irradiation-attenuated stage-3 larvae and with exported larval antigens, Experimental Parasitology *73*, 184-196.

Lustigman, S. (1993). Moulting, enzymes and new targets for chemotherapy *Onchocerca volvulus*, Parasitology Today *9*, 294-297.

Lustigman, S., Brotman, B., Huima, T., and Prince, A. M. (1991). Characterization of an *Onchocerca volvulus* cDNA clone encoding a genus specific antigen present in infective larvae and adult worms, Molecular and Biochemical Parasitology *45*, 65-75.

Lustigman, S., Brotman, B., Huima, T., Prince, A. M., and McKerrow, J. H. (1992). Molecular cloning and characterization of onchocystatin, a cysteine proteinase inhibitor of *Onchocerca volvulus*, Journal of Biological Chemistry *267*, 17339-17346.

MacDonald, S. M. (1996). Histamine-releasing factors, Current Opinion Immunology *8*, 778-783.

MacDonald, S. M. (1997). Human recombinant histamine-releasing factor, International Archieves of Allergy and Immunology *113*, 187-189.

MacDonald, S. M., Rafnar, T., Langdon, J., and Lichtenstein, L. M. (1995). Molecular identification of an IgE-dependent histamine-releasing factor, Science *269*, 688-90.

MacDonald, W. W., and Ramachandran, C. P. (1965). The influence of the gene f^m (filarial susceptibility, *Brugia malayi*) on the susceptibility of *Aedes aegypti* to seven strains of *Brugia*, *Wuchereria*, and *Dirofilaria*, Annals of Tropical Medicine and Parsitology *59*, 64-73.

Mackenzie, C. D., and Kron, M. A. (1985). Diethylcarbamazine: A review of its action in onchocerciasis, lymphatic filariasis and inflammation., Tropical Disease Bullitin *82*, R1-37.

Maekawa, Y., Himeno, K., Ishikawa, H., Hisaeda, H., Sakai, T., Dainichi, T., Asao, T., Good, R. A., and Katunuma, N. (1998). Switch of CD4+ T cell differentiation from Th2 to Th1 by treatment with cathepsin B inhibitor in experimental leishmaniasis, Journal of Immunology *161*, 2120-2127.

Mahanty, S., Ravichandran, M., Raman, U., Jayaraman, K., Kumaraswami, V., and Nutman, T. B. (1997). Regulation of parasite antigen-driven immune responses by interleukin- 10 (IL-10) and IL-12 in lymphatic filariasis, Infection and Immunity *65*, 1742-1747.

Maizels, R. M., Bundy, D. A., Selkirk, M. E., Smith, D. F., and Anderson, R. M. (1993). Immunological modulation and evasion by helminth parasites in human populations, Nature *365*, 797-805.

Maizels, R. M., Burke, J., Sutanto, I., Purnomo., and Partono, F. (1986). Secreted and surface antigens from larval stages of *Wuchereria bancrofti*, the major human lymphatic filarial parasite, Molecular and Biochemical Parasitology *19*, 27-34.

Maizels, R. M., Denham, D. A., and Sutanto, I. (1985). Secreted and circulating antigens of the filarial parasite *Brugia pahangi*: analysis of *in vitro* released components and detection of parasite products *in vivo*, Molecular and Biochemical Parasitology *17*, 277-288.

Maizels, R. M., Desavigny, D., and Ogilvie, B. M. (1984). Characterisation of surface and excretory-secretory antigens of *Toxacara canis* infective larvae, Parasite Immunology *6*, 23-37.

Maizels, R. M., Holland, M. J., Falcone, F. H., Zang, X. X., and Yazdanbakhsh, M. (1999). Vaccination against helminth parasites: the ultimate challenge for immunologists?, Immunological Reviews *171*, 125-147.

Maizels, R. M., and Lawrence, R. A. (1991). Immunological tolerance : The key feature in human filariasis ?, Parasitology Today *7*, 271-276.

Maizels, R. M., Sartono, E., Kurniawan, A., Selkirk, M. E., Partono, F., and Yazdanbakhsh, M. (1995). T cell activation and the balance of antibody isotypes in human filariasis., Parasitology Today *11*, 50-56.

Maizels, R. M., Tetteh, K. K. A., and Loukas, A. (2000). *Toxacara canis*: genes expressed by the arrested infective larval stage of parasitic nematode., International Journal of Parasitology *30*, 495-508.

Manoury, B., Hewitt, E. W., Morrice, N., Dando, P. M., Barrett, A. J., and Watts, C. (1998). An asparaginyl endopeptidase processes a microbial antigen for class II MHC presentation, Nature *396*, 695-699.

Marechal, P., Le Goff, L., Hoffman, W., Rapp, J., Oswald, I. P., Ombrouck, C., Taylor, D. W., Bain, O., and Petit, G. (1997). Immune response to the filaria *Litomosoides sigmodontis* in susceptible and resistant mice., Parasite Immunology *19*, 273-279.

Marechal, P., Le Goff, L., Petit, G., Diagne, M., Taylor, D. W., and Bain, O. (1996). The fate of the filaria *Litomosoides sigmodontis* in susceptible and resistant mice., Parasite *3*, 25-31.

Maroney, P. A., Denker, J. A., Darzynkiewicz, E., Laneve, R., and Nilsen, T. W. (1995). Most mRNAs in the nematode *Ascaris lumbricoides* are trans-spliced: a role for spliced leader addition in translational efficiency, RNA *1*, 714-723.

Marshall, E., and Howells, R. E. (1986). Turnover of the surface proteins of adult and third and fourth stage larval *Brugia pahangi.*, Molecular and Biochemical Parasitology *18*, 17-24.

Martin, S. A. M., Hunter, S. J., Thompson, F. J., and Devaney, E. (1996). Stage specific gene expression in the post-infective larvae of the filarial nematode *Brugia pahangi*, Molecular and Biochemical Parasitology *79*, 109-112.

Mason, R. W., Sol-Church, K., and Abrahamson, M. (1998). Amino acid substitutions in the N-terminal segment of cystatin C create selective protein inhibitors of lysosomal cysteine proteinases, Biochemical Journal *330*, 833-838.

McCall, J. W., Malone, J. B., Ah, H. S., and Thompson, P. E. (1973). Mongolian jirds (*Meriones unguiculatus*) infected with *Brugia pahangi* by the intraperitoneal route: a rich source of developing larvae, adult filariae and microfilariae, Journal of Parasitology *59*, 436.

McGreevy, P. B., Theis, J. H., Lavoipierre, M. M. J., and Clark, J. (1974). Studies on fiariasis III. *Dirofilaria immitis*: emergence of infective larvae from the mouthparts of *Aedes aegypti.*, Journal of Helminthology *48*.

McLaren, D. L. (1975). Micro-organisms in filarial larvae (Nematoda), Transactions of the Royal Society of Tropical Medicine and Hygiene *69*, 509-514.

McReynolds, L. A., Desimone, S. M., and Williams, S. A. (1986). Cloning and comparison of the repeated DNA sequences from the human filarial parasite *Brugia malayi* and the animal parasite *Brugia pahangi*, Proceedings of the National Academy of Sciences USA *83*, 797-801.

Meyrowitsch, D. W., Simonsen, P. E., and Makunde, W. H. (1995). A 16-year follow-up study on bancroftian filariasis in three communities of north-eastern Tanzania, Annals of Tropical Medicine and Parasitology *89*, 665-675.

Michael, E., Bundy, D. A. P., and Grenfell, B. T. (1996). Re-assessing the global prevalence and distribution of lymphatic filariasis, Parasitology *112*, 409-428.

Misaka, T., Kuroda, M., Iwabuchi, K., Abe, K., and Arai, S. (1996). Soyacystatin, a novel cysteine proteinase inhibitor in soybean, is distinct in protein structure and gene organization from other cystatins of animal and plant origin, European Journal of Biochemistry *240*, 609-614.

Misra, S., Tyagi, K., and Chatterjee, R. K. (1997). Experimental transmission of nocturnally periodic *Wuchereria bancrofti* to Indian leaf monkey (Presbytis entellus), Molecular and Biochemical Parasitology *86*, 155-157.

Moore, J., Tetley, L., and Devaney, E. (2000). Identification of abundant mRNAs from the third stage larvae of the parasitic nematode, *Ostertagia ostertagi*, Biochemical Journal *347 Pt 3*, 763-770.

Moulia-Pelat, J. P., Glaziou, P., Weil, G. J., Nguyen, L. N., Gaxotte, P., and Nicolas, L. (1995). Combination ivermectin plus diethylcarbamazine, a new effective tool for control of lymphatic filariasis, Tropical Medicine and Parasitology *46*, 9-12.

Muesch, A., Hartmann, E., Rohde, K., Rubartelli, A., Sitia, R., and Rapoport, T. A. (1990). A novel pathway for secretory proteins?, Trends in Biochemical Sciences *15*, 86-88.

Muller-Esterl, W., Fritz, H., Kellermann, J., Lottspeich, F., Machleidt, W., and Turk, V. (1985). Genealogy of mammalian cysteine proteinase inhibitors. Common evolutionary origin of stefins, cystatins and kininogens, FEBS Letters *191*, 221-226.

Naito, Y., Sasaki, M., Umemoto, T., Namikawa, I., Sakae, K., Ishihara, Y., Isomura, S., and Suzuki, I. (1995). Bactericidal effect of rat cystatin S on an oral bacterium Porphyromonas gingivalis, Comparative Biochemistry Physiology C: Pharmacology Toxicology and Endocrinology *110*, 71-75.

Nakagawa, T. Y., Brissette, W. H., Lira, P. D., Griffiths, R. J., Petrushova, N., Stock, J., McNeish, J. D., Eastman, S. E., Howard, E. D., Clarke, S. R., *et al.* (1999). Impaired invariant chain degradation and antigen presentation and diminished collagen-induced arthritis in cathepsin S null mice, Immunity *10*, 207-217.

Nayar, J. K., Bradley, T. J., Mikarts, L. L., and Knight, J. W. (1993). Intracellular development of subperiodic *Brugia malayi* influenced by mosquito thoracic muscle cells, Journal of Invertebrate Pathology *62*, 90-94.

Nelson, F. K., Greiner, D. L., Shultz, L. D., and Rajan, T. V. (1991). The immunodeficient SCID mouse as a model for human lymphatic filariasis, Journal of Experimental Medicine *173*, 659-63.

Nicolas, L., Plichart, C., Nguyen, L. N., and Moulia-Pelat, J. P. (1997). Reduction of *Wuchereria bancrofti* adult worm circulating antigen after annual treatments of diethylcarbamazine combined with ivermectin in French Polynesia, Journal of Infectious Diseases *175*, 489-492.

Nilsen, T. W. (1993). *Trans*-splicing of nematode premessenger RNA, Annual Review of Microbiology *47*, 413-440.

Oberlander, U., Adam, R., Berg, K., Seeber, F., and Lucius, R. (1995). Molecular cloning and characterization of the filarial LIM domain proteins AvL3-1 and OvL3-1, Experimental Parasitology *81*, 592-599.

Oothuman, P., Denham, D. A., McGreevy, P. B., Nelson, G. S., and Rogers, R. (1979). Successful vaccination of cats against *Brugia pahangi* with larvae attenuated by irradiation with 10 krad cobalt 60, Parasite Immunology *1*, 209-216.

Oothuman, P., Simpson, M. G., and Lawrence, B. R. (1974). Abnormal development of a filarial worm, *Brugia patei* (Buckley, Nelson and Heisch), in a mosquito host, *Anopheles labranchiae atroparvus* van Theil, Journal of Helminthology *48*, 161-165.

Osborne, J., and Devaney, E. (1998). The L3 of Brugia induces a Th2-polarized response following activation of an IL-4-producing CD4-CD8- ab T cell population, International Immunology *10*, 1583-1590.

Ottesen, E. A. (1992). The Wellcome Trust Lecture. Infection and disease in lymphatic filariasis: an immunological perspective, Parasitology *104*, S71-9.

Ottesen, E. A., Duke, B. O. L., Karam, M., and Behbehani, K. (1997). Strategies and tools for the control/elimination of lymphatic filariasis, Bulletin of the World Health Organization *75*, 491-503.

Ou, X., Tang, L., McCrossan, M., Henkle-Duhrsen, K., and Selkirk, M. E. (1995). *Brugia malayi*: localisation and differential expression of extracellular and cytoplasmic CuZn superoxide dismutases in adults and microfilariae, Experimental Parasitology *80*, 515-529.

Partono, F. (1987). The spectrum of disease in lymphatic filariasis. In Filariasis, D. Evered, and S. Clark, eds. (Chichester, John Wiley and Sons Ltd.), pp. 15-26.

Pastrana, D. V., Raghavan, N., FitzGerald, P., Eisinger, S. W., Metz, C., Bucala, R., Schleimer, R. P., Bickel, C., and Scott, A. L. (1998). Filarial nematode parasites secrete a homologue of the human cytokine macrophage migration inhibitory factor, Infection and Immunity *66*, 5955-5963.

Pay, A., Heberle-Bors, E., and Hirt, H. (1992). An alfalfa cDNA encodes a protein with homology to translationally controlled human tumor protein, Plant Molecular Biology *19*, 501-503.

Peacock, R., and Poynter, D. (1980). Field experience with a bovine lungworm vaccine. In Symposia of the British Society for Parasitology, A. E. R. Taylor, and R. Muller, eds., pp. 141-148.

Pennacchio, L. A., Lehesjoki, A. E., Stone, N. E., Willour, V. L., Virtaneva, K., Miao, J., D'Amato, E., Ramirez, L., Faham, M., Koskiniemi, M., *et al.* (1996). Mutations in the gene encoding cystatin B in progressive myoclonus epilepsy (EPM1), Science *271*, 1731-1734.

Peralta, M. E., Schmitz, K. A., and Rajan, T. V. (1999). Failure of highly immunogenic filarial proteins to provide host- protective immunity, Experimental Parasitology *91*, 334-340.

Petit, G., Diagne, M., Marechal, P., Owen, D., Taylor, D., and Bain, O. (1992). Maturation of the filaria *Litomosoides sigmodontis* in BALB/c mice; comparative susceptibility of nine other inbred strains, Annales de Parasitologie Humaine Comparrée *67*, 144-150.

Petralanda, I., Yarzabal, L., and Piessens, W. F. (1986). Studies on a filarial antigen with collagenase activity, Molecular and Biochemical Parasitology *19*, 51-59.

Philipp, M., Parkhouse, R. M. E., and Ogilvie, B. M. (1980). Changing proteins on the surface of a parasite, Nature *287*, 538-540.

Pierre, P., and Mellman, I. (1998). Developmental regulation of invariant chain proteolysis controls MHC class II trafficking in mouse dendritic cells, Cell *93*, 1135-1145.

Pogonoka, T., Oberlander, U., Marti, T., and Lucius, R. (1999). *Ancanthocheilonema viteae*: Characterisation of a molt-associated excretory/secretory18kDa protein., Experimental Parasitology *93*, 73-81.

Raghavan, N., Freedman, D. O., Fitzgerald, P. C., Unnasch, T. R., Ottesen, E. A., and Nutman, T. B. (1994). Cloning and characterization of a potentially protective chitinase-like recombinant antigen from *Wuchereria bancrofti*, Infection and Immunity *62*, 1901-1908.

Rathaur, S., Robertson, B. D., Selkirk, M. E., and Maizels, R. M. (1987). Secretory acetylcholinesterases from *Brugia malayi* adult and microfilarial parasites, Molecular and Biochemical Parasitology *26*, 257-265.

Rawlings, N. D., and Barrett, A. J. (1990). Evolution of proteins of the cystatin superfamily, Journal of Molecular Evolution Mol Evol *30*, 60-71.

Reiner, S. L., and Locksley, R. M. (1995). The regulation of immunity to *Leishmania major.*, Annual Review of Immunology *13*, 151-177.

Richer, J. K., Hunt, W. G., Sakanari, J. A., and Grieve, R. B. (1993). *Dirofilaria immitis*: Effect of fluoromethyl ketone cysteine protease inhibitors on the third- to fourth-stage molt, Experimental Parasitology *76*, 221-231.

Richer, J. K., Sakanari, J. A., Frank, G. R., and Grieve, R. B. (1992). *Dirofilaria immitis*: Proteases produced by third- and fourth-stage larvae, Experimental Parasitology *75*, 213-222.

Riese, R. J., Mitchell, R. N., Villadangos, J. A., Shi, G. P., Palmer, J. T., Karp, E. R., De Sanctis, G. T., Ploegh, H. L., and Chapman, H. A. (1998). Cathepsin S activity regulates antigen presentation and immunity, Journal of Clinical Investigation *101*, 2351-2363.

Riese, R. J., Wolf, P. R., Bromme, D., Natkin, L. R., Villadangos, J. A., Ploegh, H. L., and Chapman, H. A. (1996). Essential role for cathepsin S in MHC class II-associated invariant chain processing and peptide loading, Immunity *4*, 357-66.

Ross, L. H., Freedman, J. H., and Rubin, C. S. (1995). Structure and expression of novel spliced leader RNA genes in *Caenorhabditis elegans*, Journal of Biological Chemistry *270*, 22066-22075.

Rothstein, N., Stoller, T. J., and Rajan, T. V. (1988). DNA base composition of filarial nematodes, Parasitology *97*, 75-79.

Rubartelli, A., Cozzolino, F., Talio, M., and Sitia, R. (1990). A novel secretory pathway for interleukin-1 beta, a protein lacking a signal sequence, EMBO Journal *9*, 1503-1510.

Sanchez, J. C., Schaller, D., Ravier, F., Golaz, O., Jaccoud, S., Belet, M., Wilkins, M. R., James, R., Deshusses, J., and Hochstrasser, D. (1997). Translationally controlled tumor protein: a protein identified in several nontumoral cells including erythrocytes, Electrophoresis *18*, 150-5.

Sasa, M. (1976). Human Filariasis. A Global Survay of Epidemilogy and Control (Baltimore, University Park Press).

Schacher, J. F. (1962a). Developmental stages of *Brugia pahangi* in the final host, Journal of Parasitology *48*, 693-706.

Schacher, J. F. (1962b). Morphology of the microfilaria of *B. pahangi* and of the larval stages in the mosquito, Journal of Parasitology *48*, 679-692.

Schroeder, J. T., Lichtenstein, L. M., and MacDonald, S. M. (1996). An immunoglobulin E-dependent recombinant histamine-releasing factor induces interleukin-4 secretion from human basophils, Journal of Experimental Medicine *183*, 1265-1270.

Schroeder, J. T., Lichtenstein, L. M., and MacDonald, S. M. (1997). Recombinant histamine-releasing factor enhances IgE-dependent IL-4 and IL-13 secretion by human basophils, Journal of Immunology *159*, 447-452.

Scott, A. L., Ibrahim, M. S., and Tasmashiro, W. K. (1990). Surface-associated antigens of second, third and fourth stage larvae of *Dirofilaria immitis*, Acta Tropica *47*, 339-353.

Scott, A. L., Yenbutr, P., Eisinger, S. W., and Raghavan, N. (1995). Molecular cloning of the cuticular collagen gene *Bmcol-2* from *Brugia malayi.*, Molecular and Biochemical Parasitology *70*, 221-225.

Sebastiano, M., Lassandro, F., and Bazzicalupo, P. (1991). *cut*-1 a *Caenorhabditis elegans* gene coding for a dauer-specific noncollagenous component of the cuticle, Developmental Biology *146*, 519-530.

Selkirk, M. E., Gregory, W. F., Jenkins, R. E., and Maizels, R. M. (1993). Localization, turnover and conservation of gp15/400 in different stages of *Brugia malayi*, Parasitology *107*, 449-457.

Selkirk, M. E., Nielsen, L., Kelly, C., Partono, F., Sayers, G., and Maizels, R. M. (1989). Identification, synthesis and immunogenicity of cuticular collagens from the filarial nematodes *Brugia malayi* and *Brugia pahangi*, Molecular and Biochemical Parasitology *32*, 229-246.

Selkirk, M. E., Ruherford, P. J., Denham, D. A., Partono, F., and Maizels, R. M. (1987). Cloned antigen genes of *Brugia* filarial parasites, Biochemical Society Symposium *53*, 91-102.

Shenoy, R. K., Suma, T. K., Rajan, K., and Kumaraswami, V. (1998). Prevention of acute adenolymphangitis in brugian filariasis: comparison of the efficacy of ivermectin and diethylcarbamazine, each combined with local treatment of the affected limb, Annals of Tropical Medicine and Parasitology *92*, 587-594.

Shibuya, K., Kaji, H., Ohyama, Y., Tate, S., Kainosho, M., Inagaki, F., and Samejima, T. (1995). Significance of the highly conserved Gly-4 residue in human cystatin A, Journal of Biochemistry (Tokyo) *118*, 635-642.

Shiraishi, T., Mori, M., Tanaka, S., Sugimachi, K., and Akiyoshi, T. (1998). Identification of cystatin B in human esophageal carcinoma, using differential displays in which the gene expression is related to lymph- node metastasis, International Journal of Cancer *79*, 175-178.

Sotiropoulou, G., Anisowicz, A., and Sager, R. (1997). Identification, cloning, and characterization of cystatin M, a novel cysteine proteinase inhibitor, down-regulated in breast cancer, Journal of Biological Chemistry *272*, 903-910.

Spieth, J., Brooke, G., Kuersten, S., Lea, K., and Blumenthal, T. (1993). Operons in *C. elegans*: polycistronic mRNA precursors are processed by trans-splicing of SL2 to downstream coding regions, Cell *73*, 521-532.

Starich, T. A., Herman, R. K., and Shaw, J. E. (1993). Molecular and genetic analysis of *unc*-7, a *Caenorhabditis elegans* gene required for coordinated locomotion, Genetics *133*, 527-541.

Steinert, P. M., and Marekov, L. N. (1997). Direct evidence that involucrin is a major early isopeptide cross- linked component of the keratinocyte cornified cell envelope, Journal of Biological Chemistry *272*, 2021-2030.

Storey, D. M., and Al Mukhtar, A. S. (1982). Vaccination of jirds, *Meriones unguiculatus*, against *Litomosoides carinii* and *Brugia pahangi* using irradiated larvae of *L. carinii*, Tropenmedizin und Parasitologie *33*, 23-24.

Storey, N., and Philipp, M. (1992). *Brugia malayi*: patterns of expression of surface-associated antigens, Experimental Parasitology *74*, 57-68.

Stubbs, M. T., Laber, B., Bode, W., Huber, R., Jerala, R., Lenarcic, B., and Turk, V. (1990). The refined 2.4 Å X-ray crystal structure of recombinant human stefin B in complex with the cysteine proteinase papain: a novel type of proteinase inhibitor interaction, EMBO Journal *9*, 1939-1947.

Suswillo, R. R., Doenhoff, M. J., and Denham, D. A. (1981). Successful development of *Brugia pahangi* in T-cell deprived CBA mice, Acta Tropica *38*, 305-308.

Takagaki, Y., MacDonald, C. C., Shenk, T., and Manley, J. L. (1992). The human 64-kDa polyadenylylation factor contains a ribonucleoprotein- type RNA binding domain and unusual auxiliary motifs, Proceedings of the National Acadamy of Science U S A *89*, 1403-1407.

Takahashi, M., Tezuka, T., and Katunuma, N. (1996). Filaggrin linker segment peptide and cystatin alpha are parts of a complex of the cornified envelope of epidermis, Archives of Biochemistry and Biophysics *329*, 123-126.

Tang, L., Gounaris, K., Griffiths, C., and Selkirk, M. E. (1995). Heterologous expression and enzymatic properties of a selenium- independent glutathione peroxidase from the parasitic nematode *Brugia pahangi*, Journal of Biological Chemistry *270*, 18313-18318.

Tang, L., Ou, X., Henkle-Duhrsen, K., and Selkirk, M. E. (1994). Extracellular and cytoplasmic CuZn superoxide dismutases from *Brugia* lymphatic filarial nematode parasites, Infection and Immunity *62*, 961-967.

Taylor, M. J., Abdel-Wahab, N., Wu, Y., Jenkins, R. E., and Bianco, A. E. (1995). *Onchocerca volvulus* larval antigen, OvB20, induces partial protection in a rodent model of onchocerciasis, Infection and Immunity *63*, 4417-4422.

Taylor, M. J., Jenkins, R. E., and Bianco, A. E. (1996). Protective immunity induced by vaccination with *Onchocerca volvulus* tropomyosin in rodents, Parasite Immunology *18*, 219-225.

The *C. elegans* Sequencing Consortium (1998). Genome sequence of the nematode *C. elegans*: A platform for investigating biology, Nature *282*, 2012-2018.

The Filarial Genome Project (1999). Deep within the filarial genome: an update, Parasitology Today *15*, 219-224.

Thiele, H., Berger, M., Lenzner, C., Kuhn, H., and Thiele, B. J. (1998). Structure of the promoter and complete sequence of the gene coding for the rabbit translationally controlled tumor protein (TCTP) P23, European Journal of Biochemistry *257*, 62-68.

Tort, J., Brindley, P. J., Knox, D., Wolfe, K. H., and Dalton, J. P. (1999). Proteinases and associated genes of parasitic helminths, Advances in Parasitology *43*, 161-266.

Tsitilou, S. G., Rodakis, G. C., Alexopoulou, M., Kafatos, F. C., Ito, K., and Iatrou, K. (1983). Structural features of B family chorion sequences in the silkmoth *Bombyx mori*, and their evolutionary implications, EMBO Journal *2*, 1845-1852.

Turk, B., Turk, V., and Turk, D. (1997). Structural and functional aspects of papain-like cysteine proteinases and their protein inhibitors, Biological Chemistry Hoppe Seyler *378*, 141-150.

Turk, V., and Bode, W. (1991). The cystatins: protein inhibitors of cysteine proteinases, FEBS Lett *285*, 213-219.

Vanamail, P., Subramanian, S., Das, P. K., Pani, S. P., Rajagopalan, P. K., Bunday, D. A. P., and Grenfell, B. T. (1989). Estimation of age-specific rates of acquisition and loss of *Wuchereria bancrofti* infection, Transactions of the Royal Society of Tropical Medicine and Hygiene *83*, 689-694.

Verdot, L., Lalmanach, G., Vercruysse, V., Hartmann, S., Lucius, R., Hoebeke, J., Gauthier, F., and Vray, B. (1996). Cystatins up-regulate nitric oxide release from interferon-gamma- activated mouse peritoneal macrophages, Journal of Biological Chemistry *271*, 28077-28081.

Vickery, A. C., Albertine, K. H., Nayar, J. K., and Kwa, B. H. (1991). Histopathology of *Brugia malayi*-infected nude mice after immune- reconstitution, Acta Tropica *49*, 45-55.

Vickery, A. C., Nayar, J. K., and Albertine, K. H. (1985). Differential pathogenicity of *Brugia malayi*, *B. patei* and *B. pahangi* in immunodeficient nude mice, Acta Tropica *42*, 353-363.

Vidard, L., Rock, K. L., and Benacerraf, B. (1991). The generation of immunogenic peptides can be selectively increased or decreased by proteolytic enzyme inhibitors, Journal of Immunology *147*, 1786-1791.

von Heijne, G. (1986). A new method for predicting signal sequence cleavage sites, Nucleic Acids Research *14*, 4683-4690.

Waffenschmidt, S., Woessner, J. P., Beer, K., and Goodenough, U. W. (1993). Isodityrosine cross-linking mediates insolubilization of cell walls in *Chlamydomonas*, The Plant Cell *5*, 809-820.

Walsh, B. J., Gooley, A. A., Williams, K. L., and Breit, S. N. (1995). Identification of macrophage activation associated proteins by two- dimensional gel electrophoresis and microsequencing, Journal of Leukocyte Biology *57*, 507-512.

Wang, S. H., Zheng, H. J., Dissanayake, S., Cheng, W. F., Tao, Z. H., Lin, S. Z., and Piessens, W. F. (1997). Evaluation of recombinant chitinase and SXP1 antigens as antimicrofilarial vaccines, American Journal of Tropical Medicine and Hygiene *56*, 474-481.

Wantke, F., MacGlashan, D. W., Langdon, J. M., and MacDonald, S. M. (1999). The human recombinant histamine releasing factor: functional evidence that it does not bind to the IgE molecule, J Allergy and Clinical Immunology *103*, 642-648.

Watts, C. (1997). Capture and processing of exogenous antigens for presentation on MHC molecules, Annual Review of Immunology *15*, 821-50.

Williams, S. A., Lizotte-Waniewski, M. R., Foster, J., Guiliano, D., Daub, J., Scott, A. L., Slatko, B., and Blaxter, M. L. (2000). The filarial genome project: analysis of the nuclear, mitochondrial and endosymbiont genomes of *Brugia malayi*, International Journal of Parasitology *30*, 411-419.

Williams, S. A., The Filarial Genome Project, Johnston, D. A., and The Schistosome Genome Project (1999). Helminth genome analysis: the current status of the filarial and schistosome genome projects, Parasitology *118*, S19-S38.

Wong, M. M., Fredericks, H. J., and Ramachandran, C. P. (1969). Studies on immunization against *Brugia malayi* infection in the rhesus monkey, Bulletin of the World Health Organization *40*, 493-501.

Woolhouse, M. E. J., and Hagan, P. (1999). Seeking the ghosts of worms past, Nature Medicine *5*, 1225-1227.

World Health Orgainization (1995). World Health Orgainization Report 1995, Geneva, World Health Orgainization.

Xie, H., Bain, O., and Williams, S. A. (1994). Molecular phylogenetic studies on *Brugia* filariae using the *HhaI* repeat sequences, Parasite *1*, 255-60.

Xu, A., Bellamy, A. R., and Taylor, J. A. (1999). Expression of translationally controlled tumour protein is regulated by calcium at both the transcriptional and post-transcriptional level, Biochemical Journal *342 Pt 3*, 683-689.

Yates, J. A., and Higashi, G. I. (1985). *Brugia malayi :* vaccination of jirds with ^{60}Cobalt-attenuated infective stage larvae protects against homologous challenge., American Journal of Tropical Medicine and Hygiene *34*, 1132-1137.

Yenbutr, P., and Scott, A. L. (1995). Molecular cloning of a serine proteinase inhibitor from *Brugia malayi*, Infection and Immunity *63*, 1745-1753.

Yenofsky, R., Cereghini, S., Krowczynska, A., and Brawerman, G. (1983). Regulation of mRNA utilization in mouse erythroleukemia cells induced to differentiate by exposure to dimethyl sulfoxide, Molecular and Cellular Biology *3*, 1197-1203.

Zang, X., Yazdanbakhsh, M., Jiang, H., Kanost, M. R., and Maizels, R. M. (1999). A novel serpin expressed by blood-borne microfilariae of the parasitic nematode *Brugia malayi* inhibits human neutrophil serine proteinases, Blood *94*, 1418-1428.

Zhao, X., Ferdig, M. T., Li, J., and Christensen, B. M. (1995). Biochemical pathway of melanotic encapsulation of *Brugia malayi* in the mosquito, Armigeres subalbatus, Developmental and Comparative Immunology *19*, 205-215.

Zorio, D. A., Cheng, N. N., Blumenthal, T., and Spieth, J. (1994). Operons as a common form of chromosomal organization in *C. elegans*, Nature *372*, 270-272.